ysics Lib

VOLUME 41 SECOND SERIES NUMBER 1

Copy 2

JUL 11 1932

THE PHYSICAL REVIEW

A JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS

CONDUCTED BY

THE AMERICAN PHYSICAL SOCIETY

BOARD OF EDITORS

Managing Editor, JOHN T. TATE, University of Minnesota, Minneapolis
Assistant Editor, J. W. BUCHTA, University of Minnesota, Minneapolis
K. F. HERZFELD, A. W. HULL, F. W. LOOMIS, A. J. DEMPSTER, DAVID M. DENNISON, J. C. SLATER, R. C. GIBBS, E. O. LAWRENCE, J. H. VAN VLECK

JULY 1, 1932

Reflecting and Resolving Power of Calcite for X-rays – SAMUEL K. ALLISON 1
Scattering of X-rays by Gases and Crystals – Y. H. WOO 21
Ionization of Air by γ-Rays as a Function of Pressure and Collecting Field – I. S. BOWEN 26
Search for Isotopes of Hydrogen and Helium – WALKER BLEAKNEY 33
Elastic and Inelastic Electron Scattering in Hydrogen – A. L. HUGHES AND J. H. MCMILLEN 39
Electronic Structures of Polyatomic Molecules and Valence. II. General Considerations – ROBERT S. MULLIKEN 49
Evidence of Space Quantization of Atoms upon Impact – H. KUHN AND O. OLDENBERG 72
Equation of State of a Non-ideal Einstein-Bose or Fermi-Dirac Gas – G. E. UHLENBECK AND L. GROPPER 79
On Ferromagnetism and Related Problems of the Theory of Electrons – PAUL S. EPSTEIN 91
Errata: Shape of an X-ray Line – ARCHER HOYT 110
Letters to the Editor – 111
Vapour Pressure of Sulphur at 50°C – R. R. H. BROWN AND J. J. MUIR 111
Variation of the Cosmic Rays with Latitude – ARTHUR H. COMPTON 111
Production of Ultra-short Electromagnetic Waves – ARMIDO GRACCHEN 113
Predissociation in Nitrogen – JOSEPH KAPLAN 114
Isotopic Weight of H² – K. T. BAINBRIDGE 115
Book Reviews – 116
Current Literature of Physics

PUBLISHED BY THE AMERICAN PHYSICAL SOCIETY
[illegible]

THE AMERICAN PHYSICAL SOCIETY

OFFICERS OF THE SOCIETY

President:	W. F. G. Swann	Bartol Research Foundation, Philadelphia, Pa.
Vice-President:	Paul D. Foote	The Gulf Companies, Pittsburgh, Pa.
Secretary:	W. L. Severinghaus	Columbia University, New York City
Treasurer:	G. B. Pegram	Columbia University, New York City
Editor:	John T. Tate	University of Minnesota, Minneapolis, Minn.
Local Secretary for the Pacific Coast:	L. B. Loeb	University of California, Berkeley, Calif.

THE PHYSICAL REVIEW

Published by the American Physical Society. Editor's office: University of Minnesota, Minneapolis, Minn. Four volumes annually, six issues per volume.

First series. From July 1, 1893, to January 1, 1913, the Review was conducted by Edward L. Nichols, Ernest Merritt and Frederick Bedell, with the financial support of Cornell University. The first four volumes were issued annually; subsequent volumes, semi-annually; thirty-five volumes in all.

Second series. On January 1, 1913, the Review was transferred to the American Physical Society as announced on the first page of the new series which was then begun. Since then it has been conducted by a Board of Editors elected by the Society. Volumes 1-38 are semi-annual, beginning in January and July of each year from 1913 to 1931. Volumes 1-33 contain six issues each. Volumes 34-38 twelve issues. Beginning with Volume 39 there will be four volumes per year of six issues each.

Contributors are referred to page 2 of the advertisements for further information.

Annual subscription, throughout the Postal Union, $12.00 in advance.

General index, 1893-1920, $4.00 in advance.

Back Volumes. Volumes 1-38, $9.00 per volume; Volume 39, $4.50; Back Numbers, $1.50 per copy for issues prior to July 1, 1929 (Vols. 1-33), $.75 per copy for issues subsequent thereto. Postage is free if remittance accompanies order; otherwise postage is extra. No bound volumes can be supplied. The Review offers to purchase a limited number of First Series, Vol. 1, No. 1 (at $2.00 each); Second Series, Vol. 7, No. 1, January 1916; Vol. 7, No. 2, February 1916; Vol. 7, No. 4, April 1916; and Vol. 7, No. 5, May 1916 (at $2.00 each).

Advertisements. Rates will be supplied by the Editor on request. Copy for each number must be received at Minneapolis two weeks prior to the dates of issue; old copy is used until new copy is received. Send electrotypes direct to printers, George Banta Publishing Co., Menasha, Wis.

Discount of ten percent is allowed to members of the American Physical Society on purchase of back numbers.

Changes of Address and Complaints of failure to receive the Review, in the case of members of the Physical Society, should be addressed to the Treasurer; in the case of other subscribers, to the Editorial Secretary. New copies can be sent free in response to complaints of non-delivery only if notice of non-delivery is received within three months of date of issue.

Manuscripts for publication should be submitted to the Editor, Physical Review, University of Minnesota, Minneapolis, Minn.

Proof and all correspondence concerning papers in process of publication should be addressed to the Editorial Secretary, American Institute of Physics, 11 East 38th Street, New York, New York.

Subscriptions, renewals, and orders for back numbers should be addressed to the Editorial Secretary, American Institute of Physics, 11 East 38th Street, New York, New York.

Entered at the post office at Menasha, Wis., as second class matter.
Acceptance for mailing at special rate of postage provided for in section 1103, Act of October 3, 1917, authorized September 6, 1926.

Second Series *July 1, 1932* *Vol. 41, No. 1*

THE

PHYSICAL REVIEW

The Reflecting and Resolving Power of Calcite for X-rays

By SAMUEL K. ALLISON

University of Chicago

(Received May 20, 1932)

With a precision double spectrometer, the rocking curves from three pairs of calcites in the (1, −1) positive have been observed for seven wave-lengths from 0.21A. (W $K\alpha_1$) to 2.3A. (Cr $K\alpha_1$). Three quantities were observed at each wave-length, (1) the percent reflection (P_0), defined as the ratio of the maximum ionization current obtainable from crystal B to the ionization current produced by the x-rays incident on B from A, (2) the coefficient of reflection (R), defined as the ratio of the area of the rocking curve from B to the intensity incident on B from A, (3) the half width at half maximum of the rocking curve from B. One of the pairs (crystals III) was obviously more perfect than the others, and its values of P_0, R, and w are compared with theory. It is shown that none of the three quantities defined above is simply related to the analogous observations on a single crystal with monochromatic, parallel, incident x-rays, but that numerical predictions of the values of the above quantities may be obtained from somewhat laborious calculations based on Darwin's theory of diffraction by a perfect crystal as modified by Prins to take account of absorption. It is found that the predicted and observed coefficients of reflection are in good agreement in the wave-length range between 0.5 and 2.3A. On the basis of this agreement it is possible to state that calcite surfaces may be obtained for which there is no evidence of mosaic structure from measurements of the coefficient of reflection by double spectrometer methods in the wave-length region between 0.5 and 2.3A.

EXPERIMENTAL PART

Introduction

THE experiments reported here are similar to those performed on calcite in previous researches by other investigators.[1] The object of the experiments in the present case was primarily to provide a set of calibrated crystals to use in intensity measurements in this laboratory. The method consists essentially in comparing the intensity reflected from crystal B of a double spectrometer with that incident on B from crystal A, the instrument being set in a parallel position. In the present experiments only the (1, −1) position in the notation of Allison and Williams[1] was used.

[1] A. H. Compton, Phys. Rev. **10,** 95 (1917); Davis and Stempel, Phys. Rev. **17,** 608 (1921) and Phil. Mag. **45,** 463 (1923); Wagner and Kuhlenkampff, Ann. d. Physik **68,** 369 (1922); Allison and Williams, Phys. Rev. **35,** 1476 (1930); Davis and Purks, Phys. Rev. **34,** 181 (1929).

Apparatus

The double spectrometer used in these experiments is shown in Fig. 1. It was constructed by the Gaertner Scientific Corporation, Chicago, for the author. The principles of construction of the instrument are those previously described by Williams and Allison.[2] The advantages of this design are principally due to the fact that axis B can be rotated about axis A, and the ionization chamber can be rotated about axis B. This enables the instrument to operate both in the parallel and anti-parallel positions, and allows settings for widely different wave-lengths to be made without moving the x-ray tube. In a design for a double spectrometer recently described by A. H. Compton,[3] the x-ray tube must be rotated about axis A to change from one wave-length

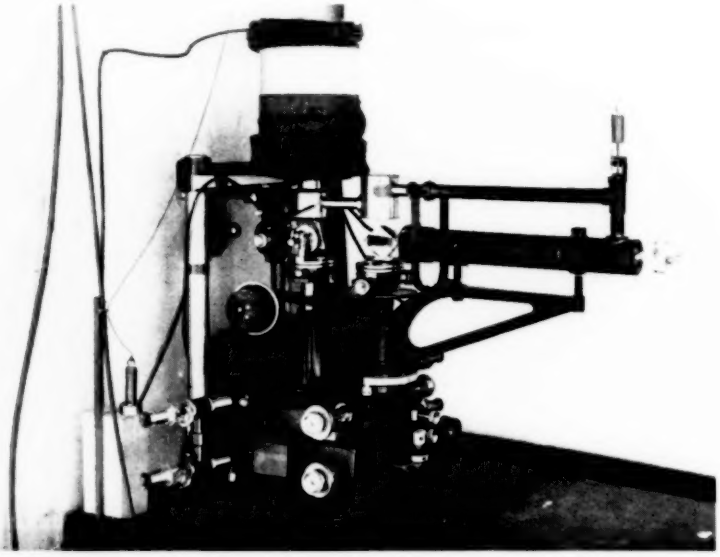

Fig. 1. The double spectrometer. The electrometer (above) is wrapped with wadding to protect it from sudden temperature changes. The instrument stands on a heavy stone slab mounted on concrete piers. The distance from the top of this slab to the shelf supporting the electrometer is 54 cm. Behind the instrument is a lead covered room in which the x-ray tubes are operated.

to another, which may be a considerable disadvantage. The author's instrument, however, is not designed for accurate *absolute* wave-length determinations, since angular measurements of rotation through large angles about axis B cannot be made with great accuracy. The present instrument is best adapted to the investigation of rocking curves, fine-structures, etc.

Although the instrument of Fig. 1 does not differ in principle from the previous one of Williams and Allison, much more care was taken in its construction. Because of the wide range of positions that axis B and the ionization chamber may assume with respect to axis A the problem of properly

[2] Williams and Allison, J.O.S.A. and R.S.I. **18,** 473 (1929).
[3] A. H. Compton, Rev. of Sci. Instr. **2,** 365 (1931).

balancing the weight on the various members is complicated. This problem was much more successfully handled in the present instrument than in its prototype. In particular, the essential requisite of the instrument namely that axis A be parallel to axis B was carefully worked out by scraping the face plates which connect the axes until as tested by a sensitive spirit level the two were parallel to less than one minute of arc.

The crystal mountings are somewhat more elaborate than those of the previous instrument in that two controlled horizontal translations are possible. In addition to the motion perpendicular to the reflecting face of the crystal, it is possible to shift the entire mounting parallel to the reflecting face. In this way the x-ray beam may be shifted along the surface of the crystal without moving that surface out of the axis of rotation. An attempt to design a plate for mounting the crystals which could readily be removed from the instrument with the crystal on it, stored in a dessicator, and then replaced, was unsuccessful. No scheme was devised such that the plate would return to its position on the axis to the required degree of accuracy (less than 10 seconds of arc) in a series of trials. The entire mounting had to be slid out of the grooved ways provided for the translatory motion in order to remove the crystals from the instrument.

Some distances on the instrument are: between the centers of the two collimating slits, 14.5 cm; from a point halfway between the two slits to axis A, 13.2 cm; from axis A to axis B, 13.8 cm; from axis B to the front window of the ionization chamber, 5.8 cm.

The ionization chamber is of the type mentioned in the previous paper by Williams and Allison,[2] and in the present experiments was filled with methyl bromide at one atmosphere pressure. The electrometer was of the Compton type, as manufactured by the Cambridge Instrument Company. It was operated at a sensitivity of 1.84 meters per volt as read on a scale 1.75 meters distant from the mirror of the instrument. It was shown by trial that a linear relationship between scale deflection and voltage on the unearthed quadrants existed at all points of the scale.

Commercial x-ray tubes were used for studies with the tungsten and chromium K series. The tube with the chromium target had a special glass window to transmit the radiation. The other wave-lengths used were obtained from tubes of the type previously described by the author.[4] The high voltage for operating the tubes was obtained from the generating plant described by Allison and Andrew.[5] In the experiments in which the $K\alpha_1$ lines of Mo, Cu, Fe, Ag, and the $L\beta_1$ line of Ir were used, the x-ray tubes were run at steady voltages between 30 and 42 k.v., higher values being used for Mo and Ag. Experiments with the Cr $K\alpha_1$ line were conducted at 38.4 k.v. and 26.8 k.v. With this radiation (wave-length 2.28A), about 86 percent was absorbed in the 58 cm air path alone, according to the absorption coefficient found by Spencer.[6] This raises the possibility that if the voltage on the tube is suffi-

[4] Allison, Phys. Rev. **30,** 245 (1927).
[5] Allison and Andrew, Phys. Rev. **38,** 441 (1931).
[6] Spencer, Phys. Rev. **38,** 1932 (1931).

ciently high, the radiation reflected at the glancing angle for Cr $K\alpha_1$ may contain little of this wave-length and be mostly second and higher orders of the general radiation. It was shown that this effect did not invalidate the results for chromium by using the instrument as a single crystal spectrometer in the usual way and examining the spectrum, and also by the fact that the rocking curves at the two voltages indicated above for Cr were in agreement. The tungsten K spectrum was obtained from a tube operated at 110 k.v.

Adjustments

The instrument was put in adjustment by the procedure previously described by Allison and Williams[1] with some minor refinements in technique. One large improvement over the methods of Allison and Williams resulted from the fact that it was found possible to use a telescope with a Gauss eyepiece in the adjustment of the crystals. This greatly increased the accuracy with which the cleavage faces could be set parallel to the axes of rotation. With the crystal holder removed, a piece of plane parallel glass, mounted on a small block equipped with levelling screws, was placed over the axis. The plane parallel plate and the telescope were adjusted until at two positions of the plate 180° apart the reflected cross-hairs coincided with those in the telescope. The axis of the telescope was then perpendicular to that of the spectrometer. The plane parallel plate was then removed, the crystal in its mounting attached to the axis, and the crystal tilted until the reflected cross-hairs from its face coincided with those in the telescope. In this way it was found possible to set the crystals accurately enough to obtain the minimum rocking curve width in parallel positions at once, and the tilting process previously described by Allison and Williams was not necessary. The telescope was placed roughly 2 meters from the spectrometer, and the crystals could be set with their faces within ± 15 seconds of arc of the axes of rotation by this method.

In agreement with previous observers the rocking curves were found to be independent of the width of the slits used to limit the horizontal divergence of the beam of x-rays. The same slit widths, however, were not used at all wave-lengths. The inadvisability of such a procedure may be realised by considering the area on the crystal face irradiated by the x-ray beam at different glancing angles. For a constant vertical divergence of the beam this area is proportional to its horizontal extent, which to a sufficient approximation is given by

$$s = 2L_c a/L \sin\theta \tag{1}$$

where L_c is the distance from a point midway between the slits to the axis of the crystal, a is the slit width (assumed the same for both slits), L is the distance between the slits, and θ is the glancing angle. In the experiments on the tungsten K radiation, $\sin\theta = 0.034$, and the irradiated area becomes large relative to that used for softer wave-lengths. The range of wave-lengths present in the beam after reflection from crystal A can be calculated from the slit width and the distance between the slits by the formula

$$d\lambda = \lambda \cot\theta d\theta = (2a\lambda \cot\theta)/L. \tag{2}$$

The various values of $d\lambda$ used are shown in column 4 of Table I. The vertical divergence of the beam was limited by a slit over the front window of the ionization chamber, at a distance of 5.8 cm from axis *B*. The other limitation was the finite size of the focal spot, certainly not greater than 0.8 cm in dia-

TABLE I. *Adjustments.*

Radiation	Crystals	*a*	$d\lambda$	*h*	Area irradiated on crystal *B*.
W $K\alpha_1$	II, V	0.012 cm	6.2 X.U.	0.8 cm	1.0 cm²
"	III	0.012	6.2	0.5	0.65
Ag $K\alpha_1$	II	0.030	26	0.2	0.24
"	III	0.040	34	0.2	0.32
"	V	0.015	13	0.8	0.48
Mo $K\alpha_1$	II	0.030	25	0.2	0.19
"	III	0.040	33	0.2	0.25
"	V	0.015	12	0.8	0.42
Ir $L\beta_1$	II	0.030	25	0.2	0.12
"	III	0.040	33	0.2	0.16
"	V	0.015	12	0.8	0.23
Cu $K\alpha_1$	II	0.030	24	0.2	0.09
"	III	0.040	32	0.2	0.12
"	V	0.020	16	0.8	0.23
Fe $K\alpha_1$	II	0.030	24	0.2	0.07
"	III	0.040	32	0.2	0.09
"	V	0.020	16	0.8	0.18
Cr $K\alpha_1$	II, III, V	0.030	23	0.8	0.24

eter and 58 cm distant from this slit. The width of the slit in front of the ionization chamber limiting the vertical divergence of the beam in various experiments is given as *h* in column 5 of Table I. In column 6 of the same table is given the area irradiated on crystal *B* in cm², obtained roughly by multiplying the values in column 5 by the appropriate value of *s* obtained from Eq. (1).

Crystals

Five pairs of split calcites were examined in this work. Two of the pairs, calcites I and IV, were examined only for a few wave-lengths and the data are not reported here. The three pairs II, III, and V had freshly cleaved surfaces. They were clear calcites, suitable for the construction of Nicol prisms, but of unknown origin. They were split by a method recommended to the author by Professor Bergen Davis. The sample to be split was mounted on the carriage of a milling machine, in such a way that the lateral motion of the carriage was accurately parallel to the cleavage direction. A very thin circular saw was then run at slow speed and a groove made across the specimen. By continued sawing in this groove a split soon started and the slight jars due to the impact of the saw teeth on the crystal spread it until fissure was complete. The two halves of the crystal were then immediately placed in a dessicator and kept until used. The crystals were parallelopipeds usually from 0.5 to 1 cm thick and some 5 cm on a side.

The ability of the crystal surfaces to give a sharp reflected image of the cross hairs of the Gauss eye-piece was different in the various specimens. It is a rather disconcerting fact that crystals III, which proved to give the high-

est resolving power for x-rays, were the poorest optical reflectors. Crystals II gave a very clear reflected image, as good as that from a polished glass plate. Crystals V were intermediate between II and III in this respect.

The most extensive investigations were carried out on crystals III, and before these crystals were finally removed from the instrument a test was performed to make sure that the resolving power had not changed during the experiments. On Nov. 19, 1931 these crystals gave a rocking curve of 2.6″ half width at half maximum for Mo $K\alpha_1(1, -1)$. After the investigation of the other wave-lengths was completed, on Jan. 2, 1932, the experiment with Mo $K\alpha_1$ was repeated and again 2.6″ obtained. There is therefore no evidence that the crystals III had changed during the period of experimentation.

Measurements

From the rocking curves for a given wave-length three quantities were determined. The half width at half maximum, w, is the angular range in seconds of arc through which crystal B must be turned to reduce the power entering the ionization chamber to one-half its maximum value. By maximum value is meant the largest ionization current observed, corrected for base-line. This base-line was taken as the ionization current observed when crystal B was rotated from its angular position for maximum reflection through an angular displacement of 6 or more times the half width at half maximum. The percent reflection, P_0, is the ratio of the maximum ionization current obtainable by reflection from crystal B to that obtainable from the beam incident on B from A. The coefficient of reflection, R, is in this paper defined as the ratio of the area under the rocking curve to the ionization current obtainable from the beam incident on crystal B from crystal A. Values of R observed and calculated in this paper are based on the radian as the unit of angle.

A regular routine was developed for the recording of rocking curves, the steps of which will now be described. A run was begun with crystal B out of the x-ray beam. The beam reflected from crystal A (which had been set at the proper angle) was allowed to enter the ionization chamber, and the current through the x-ray tube was adjusted until the rate of deflection of the electrometer was between 5 and 10 scale divisions per second. When conditions had become steady, a series of readings, usually three, were made on the beam reflected from crystal A. Between these readings the ionization chamber was shifted slightly to make perfectly sure that the entire beam was entering the window. If these three readings agreed to within 1 percent, the ionization chamber was displaced 5° and a reading taken of the diffusely scattered radiation. The difference between the average of the three readings and the base-line reading was taken as proportional to the intensity incident on B from A. Except in experiments on W $K\alpha_1$ the base-line was negligible here. Crystal B was then advanced into the beam, and the ionization chamber placed to receive the reflected beam from it. The rocking curve was then taken. Care was taken to obtain readings far enough from the maximum to supply a reliable base. Crystal B was then removed, and the power incident upon it again determined; the average of the initial and final values being used as the inten-

sity incident on B during the taking of the rocking curve. The results are given in Table II. Each value is the result of at least two independent observations.

TABLE II. *Experimental results.*
w = half width at half maximum of rocking curve (seconds).
P_0 = percent reflection.
R = coefficient of reflection, in (radians)$^{-1}$.

Line	λ (X.U.)	w II	w III	w V	$100\,P_0$ II	$100\,P_0$ III	$100\,P_0$ V	$R\times10^5$ II	$R\times10^5$ III	$R\times10^5$ V
W $K\alpha_1$	208.6	4.2	2.3	7.2	33	35	28	2.34	1.17	2.14
Ag $K\alpha_1$	558.3	3.8	2.2	7.0	48	59	34	2.15	1.53	2.38
Mo $K\alpha_1$	707.8	3.9	2.6	7.2	52	63	33	2.31	1.86	2.44
Ir $L\beta_1$	1155	4.8	4.1	8.1	59	64	37	3.49	2.95	3.24
Cu $K\alpha_1$	1537	5.6	4.9	7.8	61	62	46	3.92	3.46	3.91
Fe $K\alpha_1$	1932	6.8	6.2	9.9	54	58	42	4.45	4.18	4.45
Cr $K\alpha_1$	2285	7.7	7.1	9.3	52	55	42	4.73	4.68	4.37

If we examine the data on w in Table II, we see that the w values for crystals III are consistently lower than for crystals II or V. This is taken to indicate that crystals III approach the ideal of a perfect crystal more nearly than do the others investigated, since any mosaic structure would enlarge the angular range over which reflected can take place. The remainder of this report will be concerned with a comparison of the values of w, P_0, and R observed for crystals III with the predictions of the theory of diffraction by a perfect crystal.

THEORETICAL PART

Theory of diffraction of x-rays by an absorbing, perfect crystal

The theory of the diffraction of x-rays from the face of a perfect crystal has been developed by Darwin,[7] and later, Ewald.[8] In the discussion here we shall consider the theory as derived by Darwin. In Darwin's derivation, the absorption by incoherent processes of the x-rays as they penetrate the crystal, is neglected. The absorption due to ejection of photoelectrons or recoil electrons would be considered as an incoherent process in this sense. The result of Darwin's calculations can be expressed in the following form. The ratio of the intensity reflected from the crystal to that incident upon it is

$$\frac{I_{(\beta-\theta)}}{I_0} = \left| \frac{-q}{v + (v^2 - q^2)^{1/2}} \right|^2, \tag{3}$$

where

$$v = \left\{ 2\pi d(\beta - \theta) \cos\theta_0 \right\} / \lambda, \tag{4}$$

$$q = -\frac{2\pi d}{\lambda \sin\theta_0}\,\delta\,\frac{F(2\theta_0)}{F(0)}, \tag{5}$$

$$\theta = \theta_0 + \delta \sec\theta_0 \operatorname{cosec}\theta_0, \tag{6}$$

$$\theta_0 = \sin^{-1}\frac{n\lambda}{2d}. \tag{7}$$

[7] Darwin, Phil. Mag. **27,** 325 and 675 (1914).
[8] Ewald, Phys. Zeits. **26,** 29 (1925).

In the above expressions, $(\beta-\theta)$ is the deviation of the glancing angle from the angle θ, which is the corrected Bragg angle. λ is the wave-length of the x-rays in air; d, the grating space of the crystal; n, the order of the spectrum, and δ the deviation of the index of refraction from unity. The sign to be used for the radical in the denominator of Eq. (3) is determined by the physical requirement that the ratio on the left hand side of this equation must be less

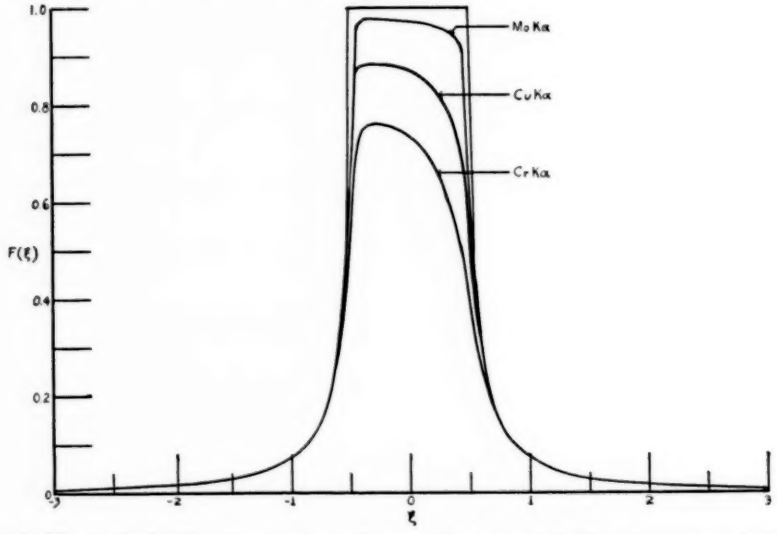

Fig. 2. Theoretical diffraction patterns from a single crystal. The outer curve is Darwin's theory, uncorrected for absorption, the inner curves illustrate Prins' absorption correction. With values of ξ plotted as abscissae, the curves apparently have about the same width at half maximum. If instead of ξ, we consider the actual angular deviations from θ, the width increases with wave-length.

than or equal to one. $F(2\theta_0)$ is the ratio of the amplitude of the wave scattered at angle $2\theta_0$ from the unit cell to the amplitude incident upon it. $F(0)$ is analogous ratio for the wave scattered in the forward direction of the incident beam. Let us first consider the simplified case in which the entire incident radiation is plane polarized in such a direction that the electric vector lies in a plane perpendicular to the plane of incidence of the radiation on the crystals. For the case,

$$F(2\theta_0)/F(0) = FZ^{-1} \tag{8}$$

where F is the structure factor and Z the number of electrons in the unit cell of the crystal. In the range $-q<v<q$, Eq. (3) gives rise to 100 percent reflection. If we call this range of angle $\Delta\theta_\sigma$ we have

$$\Delta\theta_\sigma = 4\delta FZ^{-1} \operatorname{cosec} 2\theta_0. \tag{9}$$

In this and following equations, the subscript σ indicates that the electric vector of the radiation is perpendicular to the plane of incidence. We shall use this $\Delta\theta_\sigma$ as a unit of measure of angular deviation from the corrected Bragg angle θ, and introduce the quantity ξ_σ defined by

$$\xi_\sigma = (\beta - \theta)/\Delta\theta_\sigma. \tag{10}$$

If this ξ_σ is introduced into Eq. (3) ,we obtain

$$F(\xi_\sigma) = \left| 2\xi_\sigma + (4\xi_\sigma^2 - 1)^{1/2} \right|^{-2} \tag{11}$$

A plot of $F(\xi_\sigma)$ is shown as the outer curve of Fig. 2.

Prins[9] has recently modified Darwin's theory in an attempt to take account of the absorption of the x-rays in the crystal by incoherent processes, as mentioned above. It is well known that in the classical dispersion theory, the index of refraction is the real part of a complex number, which number we may represent by $(1-\delta-i\beta)$. The real part of this number, $(1-\delta)$, is the ordinary index of refraction, i.e., the ratio of the phase velocity in vacuum to that in the medium. The imaginary part is related to the linear absorption coefficient μ_l of the radiation in the medium by

$$\beta = \lambda\mu_l/4\pi. \tag{12}$$

The essential feature of Prins' method is that he has treated the absorption in the classical manner, as if it took place by coherent processes, which in general leads to the substitution of $(\delta+i\beta)$ for δ when it appears in Darwin's treatment.

We shall take the result of Prins' calculations as it appears in Eq. (11) of his paper. This is, with a slight change of notation,

$$\frac{I_{(\beta-\theta_0)}}{I_0} = \left| \frac{d + ib}{(\beta-\theta_0)\sin\theta_0\cos\theta_0 - (\delta+i\beta) + (\{(\beta-\theta_0)\sin\theta_0\cos\theta_0 - (\delta+i\beta)\}^2 - (d+ib)^2)^{1/2}} \right|^2 \tag{13}$$

We shall first consider the application of this formula to a crystal composed of atoms of one kind only, in equally spaced planes. Under these conditions the quantity $(d+ib)$ is related to the quantity $(\delta+i\beta)$ as the amplitude of the wave scattered by an atom at angle $2\theta_0$ to the amplitude scattered in the forward direction. In order to be more specific about the relationship between the two quantities, we will adopt the simpler of Prins' two hypotheses, hypothesis *I*. This states that all electrons in the atom may be treated in the same way as regards the amplitude scattered as a function of the scattering angle, irrespective of the level to which they belong. This hypothesis leads to the relation

$$d + ib = (\delta + i\beta)f(2\theta_0)/f(0) \tag{14}$$

where $f(2\theta_0)$ is the amplitude scattered by the atom at the angle $2\theta_0$ from an incident wave of unit amplitude. A more detailed and correct procedure would be to consider the *K*, *L*, *M*, etc. electrons in the atom separately, both with respect to their contributions to δ and β, and the angular distribution of the amplitude scattered from them. In this paper, however, the simplified procedure of Eq. (14) will be adopted.

[9] Prins, Zeits. f. Physik **63**, 477 (1930).

In a crystal such as calcite, which consists of interlaced planes of different kinds of atoms; instead of $(d+ib)$ in Eq. (13), we must introduce a quantity

$$D + iB = \sum_i (\delta_i + i\beta_i) \frac{f_i(2\theta_0)}{f_i(0)} e^{2\pi n i(hx_i + ky_i + lz_i)}. \tag{15}$$

In this expression, δ_i is the contribution to δ for the medium from atom i. n is the order of the reflection; h, k, l, the Miller indices of the reflecting plane, and x_i, y_i, z_i the coordinates of the atom i in the unit cell. The summation is to be carried out over all the atoms present in the unit cell. In more common notation, for the σ type of polarization defined above,

$$f_i(2\theta_0)/f_i(0) = F_i Z_i^{-1} \tag{16}$$

where F_i is the atomic structure factor of the atom of type i, and Z_i is the number of electrons in the atom.

In calculating the values of $D+iB$ by Eq. (15), the author has made a further approximation, namely, that the contributions to δ and β from a given atom i are proportional to the fraction of the total number of electrons in the unit cell contained in the atom in question. If Z is the total number of electrons in the unit cell, this means that

$$\delta_i + i\beta_i = Z_i(\delta + i\beta)/Z. \tag{17}$$

It is clear that more accurate calculations could be made, the best method being perhaps to subdivide the electrons in the unit cell into groups according to their binding energy, but in view of the approximate nature of the graphical calculations to come later it would seem that the somewhat crude assumption of Eq. (17) is permissible.

Substitution of Eqs. (17) and (16) in Eq. (15) gives

$$D + iB = Z^{-1}(\delta + i\beta) \sum_i F_i e^{2\pi n i(hx_i + ky_i + lz_i)}. \tag{18}$$

The summation in Eq. (18) is commonly known as the structure factor F, so that finally,

$$D + iB = Z^{-1}F(\delta + i\beta). \tag{19}$$

We are now ready to adapt Eq. (13) to a crystal of more than one kind of atom, and to put it into a form similar to Eq. (11). It should be noted that in Eq. (13) the reference angle is θ_0, the uncorrected Bragg angle, whereas in Eq. (10) the reference angle is θ. In terms of θ_0, the expression for ξ_σ in Eq. (10) becomes, after making use of Eq. (6).

$$\xi_\sigma = \frac{(\beta - \theta_0)\sin\theta_0\cos\theta_0}{2\delta FZ^{-1}} - \frac{1}{2FZ^{-1}}. \tag{20}$$

Putting $D+iB$ from Eq. (19) in place of $d+ib$ in Eq. (13), and introducing ξ_σ from Eq. (20), we obtain

$$F(\xi_\sigma) = \left| \frac{FZ^{-1}(1 + i\beta/\delta)}{2\xi_\sigma FZ^{-1} - i\beta/\delta + \left[(2\xi_\sigma FZ^{-1} - i\beta/\delta)^2 - F^2Z^{-2}(1 + i\beta/\delta)^2\right]^{1/2}} \right|^2. \tag{21}$$

If the effect of absorption is neglected, that is, $\beta = 0$, this expression reduces to Eq. (11), Darwin's original form.

The calculation of the shape of the diffraction pattern to be expected from a perfect calcite crystal irradiated at the proper angles by monochromatic, plane, x-rays of wave-lengths 0.708A, (Mo $K\alpha_1$), 1.537A (Cu $K\alpha_1$), and 2.285A (Cr $K\alpha_1$) has been carried out from Eq. (21). The values of δ were obtained from Larsson.[10] The values of β may be obtained from the linear absorption coefficient by Eq. (12). The value of μ_l was calculated for Mo $K\alpha_1$ in calcite from data collected by A. H. Compton,[11] and extended to other wave-lengths by assuming the validity of the λ^3 law. The values used are shown in Table III.

TABLE III. *Values of δ and β used in Eq. (21).*

λ	δ	β	β/δ
0.708A	1.87×10^{-6}	1.31×10^{-8}	0.007
1.537	8.80	2.92×10^{-7}	0.033
2.285	19.5	1.43×10^{-6}	0.073

The value of FZ^{-1} for calcite in the first order from the cleavage faces, based on the atomic structure factors of Bragg and West[12] has been shown by Allison and Williams[1] to be 0.47. All the calculations of this paper have been based on this value, which may be somewhat low, as discussed later. Some of the values of $F(\xi_\sigma)$ for these three cases, both from Eq. (11) (Darwin) and Eq. (21) (Prins) are shown in Table IV.[13]

TABLE IV *Diffraction pattern of a perfect calcite crystal.*

Eq. (10) ξ_σ	$F(\xi_\sigma)$ Darwin	$F(\xi_\sigma)$ Prins Mo $K\alpha_1$	Cu $K\alpha_1$	Cr $K\alpha_1$
−2.13	0.014	0.014	0.014	0.014
−1.06	0.063	0.062	0.062	0.062
−0.85	0.106	0.105	0.105	0.105
−0.64	0.234	0.233	0.232	0.229
−0.43	1.000	0.959	0.874	0.709
−0.21	1.000	0.973	0.881	0.758
±0.00	1.000	0.971	0.870	0.735
+0.21	1.000	0.960	0.830	0.670
0.43	1.000	0.924	0.702	0.503
0.53	0.501	0.486	0.426	0.341
0.64	0.234	0.233	0.229	0.213
0.85	0.106	0.105	0.105	0.103
1.06	0.063	0.062	0.062	0.062
2.13	0.014	0.014	0.014	0.014

The calculated values of Table IV are shown graphically in Fig. 2. An interesting feature of the curves corrected for absorption by Prins' method is that

[10] Larsson, Diss. Upsala 1929. Reported in Siegbahn, Spektroskopie der Rontgenstrahlen, 2nd Ed., (1931) Table XIII, p. 39.

[11] A. H. Compton, X-rays and Electrons, (1926), Table VI–2 p. 182.

[12] Bragg and West, Zeits. f. Krist. **69**, 118 (1928).

[13] A large part of the calculations involved in this paper were made for me by Mrs. G. S. Monk, computer of this laboratory.

they are asymmetrical. The possibility of detecting such a lack of symmetry by double spectrometer measurements will be discussed later.

The areas under the curves of Fig. 2 are important in the theory of intensity of reflection. By integration of Eq. (11) it may be shown that

$$\int_{-\infty}^{\infty} F(\xi_\sigma) d\xi_\sigma = 4/3. \tag{22}$$

The unit of angular deviation in Eq. (22) is the angle $\Delta\theta_\sigma$, as is shown by Eq. (10). If we adopt more fundamental units, i.e., radians, the area under the diffraction pattern is $4\Delta\theta_\sigma/3$. The areas under the other curves of Fig. 2 were obtained by use of the integraph, and are shown in Table V.[14]

TABLE V. *Values of* $\int_{-\infty}^{\infty} F(\xi_\sigma) d\xi_\sigma$ *from Eqs. (11) and (21).*

Line	Darwin	Prins
Mo $K\alpha_1$	1.333	1.250
Cu $K\alpha_1$	1.333	1.122
Cr $K\alpha_1$	1.333	0.935

Calculation of the shape of the (1, −1) rocking curve for plane polarized radiation

A careful analysis of the operation of the double crystal spectrometer shows that the shape of the rocking curve obtained in parallel positions in the first order could not possibly be that of the curves in Fig. 2, even if perfect crystals were available and the instrument were in correct adjustment. Neither can be the area under a rocking curve be directly compared with the areas under the curves in Fig. 2. The fundamental physical reason for this is that all the theoretical curves previously described were calculated for parallel, monochromatic radiation. Due to the fact that the dispersion is zero in parallel positions, the spectral distribution of the radiation is not important, but, although the first crystal greatly reduces the horizontal divergence of the beam, it does not reduce it to zero, and the beam incident on crystal B is still a divergent one. The author has previously pointed out[15] that the function governing the shape of the rocking curves is

$$\Phi(\xi_{B\sigma}) = \int_{-\infty}^{\infty} F(\xi_\sigma) F(\xi_\sigma - \xi_{B\sigma}) d\xi_\sigma. \tag{23}$$

In this expression,

$$\xi_{B\sigma} = (\beta_B - \theta_B)/\Delta\theta_\sigma \tag{24}$$

[14] It is these areas which, expressed in radian measure, are usually defined as the coefficients of reflection, and it is true that the reflecting power of a single crystal is proportional to them. It is one of the principal objects of this paper to make clear that there is a discrepancy between this definition and the experimental attempts to measure this quantity by the double spectrometer. In this paper, the term coefficient of reflection refers to area under the rocking curve of crystal B. This is quite a different thing physically from the areas in Table V.

[15] S. K. Allison, Phys. Rev. **38,** 203 (1931) Eq. (8).

where $(\beta_B - \theta_B)$ is the angular deviation of crystal B from the position in which its reflecting face is parallel to that of A. $\Phi(\xi_{B\sigma})$ is the intensity of the beam entering the ionization chamber from a single position of crystal B. $F(\xi_\sigma)$ is the diffraction pattern for a single crystal as in Fig. 2. The derivation of Eq. (23) is perhaps made clearer by Fig. 3, in which the course of a ray through the instrument is shown.

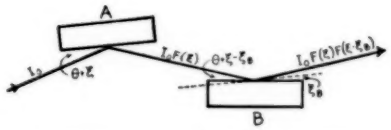

Fig. 3. Course of a ray through the double spectrometer in a parallel position, illustrating Eq. (23).

It is the custom to express the intensity reflected from B in terms of the intensity of the beam incident on B from A, which is proportional to the area under the single crystal diffraction pattern. We thus obtain the following expression

$$P_\sigma = \frac{\int_{-\infty}^{\infty} F(\xi_\sigma)F(\xi_\sigma - \xi_{B\sigma})d\xi_\sigma}{\int_{-\infty}^{\infty} F(\xi_\sigma)d\xi_\sigma} \tag{25}$$

where P_σ is the ordinate of the rocking curve from crystal B according to the convention above. The percent reflection, P_0, will be given by the value of P_σ when $\xi_{B\sigma} = 0$, or

$$P_0 = \frac{\int_{-\infty}^{\infty} [F(\xi_\sigma)]^2 d\xi_\sigma}{\int_{-\infty}^{\infty} F(\xi_\sigma)d\xi_\sigma}. \tag{26}$$

Laue[16] has discussed the possibility of solving Eq. (23), that is, of determining the shape of the single crystal diffraction pattern from the experimentally determined $\Phi(\xi_{B\sigma})$, and has come to some important conclusions. In the first place, the observed rocking curve ($\Phi(\xi_{B\sigma})$ or P_σ) will necessarily be an even function, whether or not the single crystal diffraction curves $F(\xi_\sigma)$ are asymmetrical. If we assume that the single crystal diffraction pattern is an even function, then it is possible to find its shape by solution of Eq. (23), otherwise, not. This means that by experiments of the type reported here it is impossible to discover whether the single crystal diffraction patterns are really asymmetrical as in the theory of Prins. If asymmetrical rocking curves are experimentally found, the only legitimate interpretation is that the diffraction patterns of the two crystals used are not identical. It should be remembered that the preceding statement applies only to parallel positions.

[16] Laue, Zeits. f. Physik **72**, 472 (1931).

In view of the findings of Laue, as reported above, the author has not attempted to solve the Eq. (23), using his observed $\Phi(\xi_{B\sigma})$, and discover the nature of the single crystal curve. The procedure adopted has been the reverse of this, namely to calculate $\Phi(\xi_{B\sigma})$ from Eq. (23), using theoretical $F(\xi_\sigma)$'s from Eqs. (11) and (21), and comparing the result with experiment. Calculated values of $\Phi(\xi_{B\sigma})$ are shown in Table VI. These values were obtained mainly by graphical integration.

TABLE VI. *Calculated values of* $\Phi(\xi_{B\sigma})$ *from Eq. (23). The unit of angle is* $\Delta\theta_\sigma$ *from Eq. (9).*

$\xi_{B\sigma}$	Darwin	Prins Mo $K\alpha_1$	Prins Cu $K\alpha_1$	Prins Cr $K\alpha_1$
0	1.068	0.970	0.765	0.521
±0.21	0.954	.876	.694	.475
± .43	.786	.727	.566	.401
± .64	.590	.552	.434	.308
± .85	.396	.378	.272	.212
±1.06	.209	.200	.161	.135
±1.60	.077	.070	.063	.052
±2.13	.040	.036	.033	.029
±2.66	.024	.022	.020	
±3.19	.019	.014	.013	.012

Calculation of the shape of the rocking curve with unpolarized incident radiation

Before comparing the calculations summarized in Table VI directly with experiment, we must consider the fact that after reflection from crystal A, the beam incident on crystal B is partially polarized. Since the radiation used was in each case a characteristic emission line of the target, it is known that the radiation incident on crystal A was completely unpolarized. We may then arbitrarily divide up the beam incident on A into two components of equal intensity, one polarized so that its electric vector is perpendicular to the plane of incidence (subscript σ) and one with its electric vector parallel to the plane of incidence (subscript π). We then treat each component separately in its passage through the apparatus. In Darwin's treatment, the extent of the region of 100 percent reflection with π-polarization is

$$\Delta\theta_\pi = 4\delta FZ^{-1} \operatorname{cosec} 2\theta_0 \cos 2\theta_0 = 4\delta FZ^{-1} \cot 2\theta_0 \tag{27}$$

which is to be compared with Eq. (9). We now introduce the quantities

$$\xi_\pi = (\beta - \theta)/\Delta\theta_\pi = \xi_\sigma \sec 2\theta_0 \tag{28}$$

and

$$\xi_{B\pi} = (\beta_B - \theta_B)/\Delta\theta_\pi = \xi_{B\sigma} \sec 2\theta_0. \tag{29}$$

By use of the preceding three equations, the derivation of the rocking curves for radiation polarized so that its electric vector lies in the plane of incidence (π-polarization) may be accomplished simply by changing the subscript σ to π in all the equations in which it appears up to and including Eq. (26). In adding the contributions from σ- and π-components to obtain the unpolarized values, it must be remembered that the angular unit in which σ

and π quantities are expressed is different. (Compare Eqs. (27) and (9).) We shall adopt the angular unit $\Delta\theta_\sigma$ in the following equations dealing with unpolarized radiation, and change quantities expressed in terms of $\Delta\theta_\pi$ to their values in terms of $\Delta\theta_\sigma$. We then obtain

$$P = \frac{\Phi(\xi_{B\sigma}) + \Phi(\xi_{B\pi}) \cos 2\theta_0}{\int_{-\infty}^{\infty} F(\xi_\sigma)d\xi_\sigma + \int_{-\infty}^{\infty} F(\xi_\pi)d\xi_\pi} = \frac{\Phi(\xi_{B\sigma}) + (\cos 2\theta_0)\Phi(\xi_{B\sigma} \sec 2\theta_0)}{(1 + \cos 2\theta_0) \int_{-\infty}^{\infty} F(\xi_\sigma)d\xi_\sigma} \quad (30)$$

where P is the ordinate of the rocking curve, that is, the ratio of the ionization current obtained from a setting of crystal B to the current obtained from the

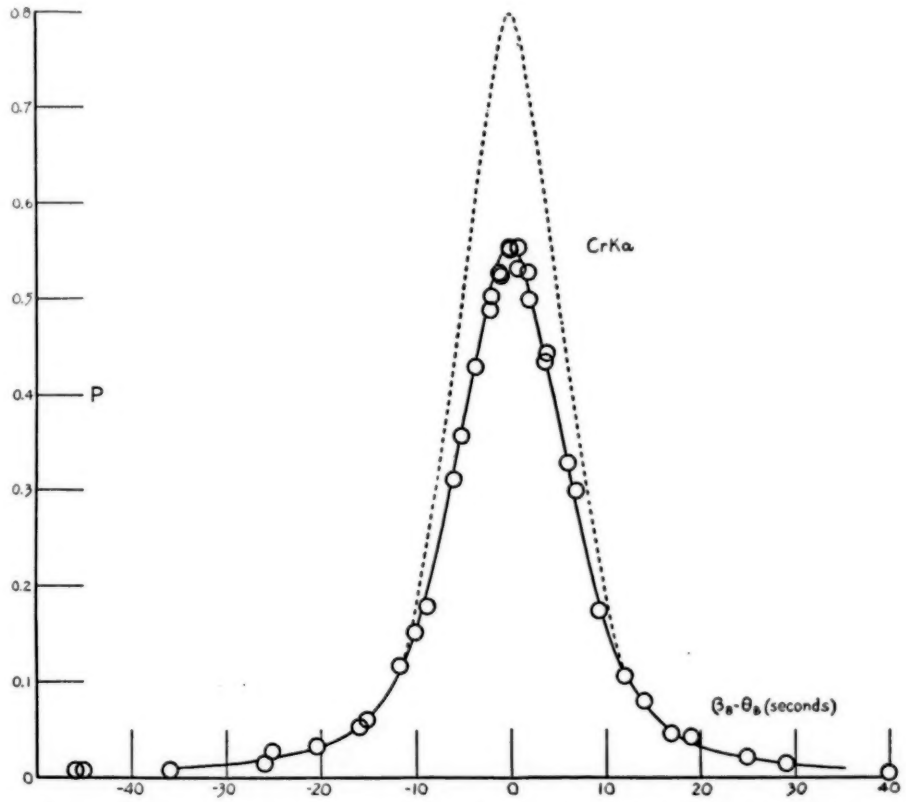

Fig. 4. Comparison of observed and calculated rocking curves for Cr $K\alpha_1$. The dotted curve is the rocking curve calculated from Darwin's theory, in which absorption in the crystal is neglected. The solid curve is the curve derived from Prins' theory. The circles represent experimental points. The points are not fitted to the curve in any way, for instance by an adjustable constant.

radiation incident on B from A. If $\xi_{B\sigma} = 0$ in Eq. (30), that is, the two crystals are accurately parallel, it reduces to Eq. (26), which means that the percent reflection, P_0, is independent of the polarization. This, of course, is not true of the coefficient of reflection, R.

The use of Eq. (30) will be illustrated by a sample calculation of a point on

the rocking curve for Cr $K\alpha_1$ (1, −1) according to Prins. The denominator of Eq. (30) can be obtained from a knowledge of the Bragg angle for Cr $K\alpha_1$ on calcite, with the values in Table V. Its value is (1+0.715) (0.935) = 1.603. Let us consider a setting of crystal B such that $\xi_{B\sigma} = 0.64$. From Table VI we find $\Phi(\xi_{B\sigma}) = 0.308$. The value of $\xi_{B\sigma}$ sec $2\theta_0$ is $0.64 \times 1.40 = 0.896$. From interpolation of Table VI, we find $\Phi(0.896) = 0.194$, and $\Phi(0.896)\cos 2\theta$ is 0.139. Thus P for this setting of the crystal is 0.279.

The calculated values of $\Delta\theta_\sigma$ from Eq. (9) are shown in Table VIII. They are computed for $FZ^{-1} = 0.47$, using δ-values from Table III, with the exception of W $K\alpha_1$, where δ is extrapolated by assuming a variation with λ^2. The rocking curves to be expected from calcite for Cr $K\alpha_1$ (1, −1) with unpolarized radiation from Darwin's and Prins equations can be plotted from the calculated values shown in Table VII.

TABLE VII. *Theoretical form of the calcite rocking curve* Cr $K\alpha_1$ (1, −1).

$\xi_{B\sigma}$	$\beta_B - \theta_B$ (seconds)	P Darwin	P Prins
0.0	0	0.798	0.557
± .21	2.27	.698	.496
± .43	4.64	.540	.396
± .64	6.91	.370	.279
± .85	9.18	.229	.183
±1.06	11.4	.125	.116
±1.60	17.3	.045	.045
±2.13	23.0	.024	.025
±2.66	28.7	.015	.016
±3.19	34.5	.011	.010

TABLE VIII. *Values of* $\Delta\theta_\sigma$.

Line	λ	cosec 2θ	$\Delta\theta_\sigma$ radians	$\Delta\theta_\sigma$ seconds
W $K\alpha_1$	0.2086A	14.44	4.4×10^{-6}	0.91
Mo $K\alpha_1$	0.7078	4.31	15.2	3.14
Cu $K\alpha_1$	1.537	2.04	33.7	6.96
Cr $K\alpha_1$	2.285	1.43	52.4	10.8

The calculated rocking curves and the experimental points for Cr $K\alpha_1$ taken on crystals III are shown in Fig. 4. It is seen that the calculated rocking curve agrees within the experimental error with that obtained experimentally. The half width at half maximum of the theoretical curve is 6.9″, the observed value is 7.1″. The percent reflection P_0 is 0.557 from Table VII; the observed value is 0.554. It is of interest to note that according to the analysis of the action of the double spectrometer presented in this paper, even if the diffraction pattern from each of the two crystals could be treated by Darwin's equation, the observed percent reflection would only be 0.798 as against 1.00 which has been expected by some writers.

The shape of the rocking curve to be expected from Prins' theory has been calculated for Mo $K\alpha_1$, Cu $K\alpha_1$ and Cr $K\alpha_1$. The rocking curve to be expected from W $K\alpha_1$ has been calculated from Darwin's equation directly. The re-

sultant P_0's and half widths at half maximum, w, are compared with experiment in Fig. 5. The resolving power is calculated from a previous definition of the author,[17] which is

$$\lambda/d\lambda = D\lambda/2w. \tag{31}$$

This is to be interpreted as the resolving power which the instrument would have in the (1, 1) position with the same crystals. D is the dispersion of the instrument in the (1, 1) position.

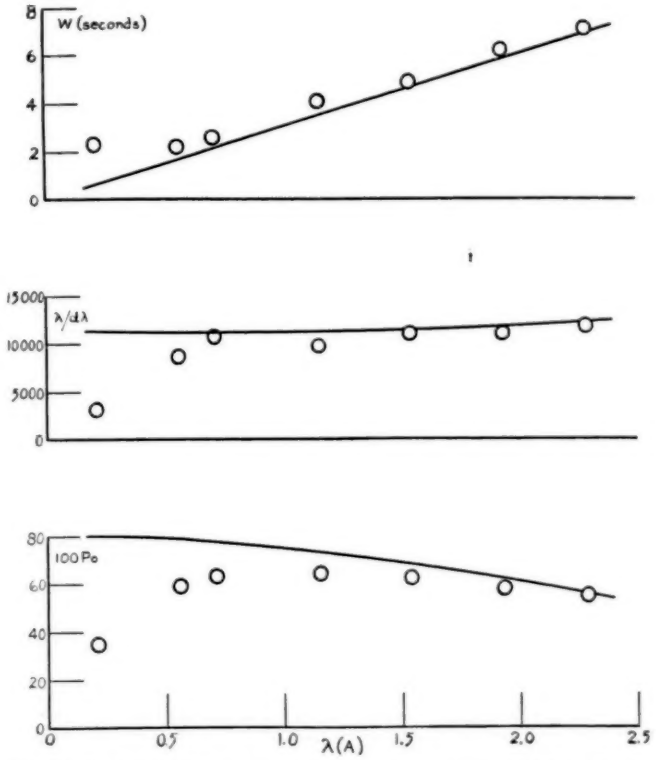

Fig. 5. The half width at half maximum (w), resolving power ($\lambda/d\lambda$), and percent reflection (P_0) of calcites III compared with theoretical values deduced from Prins. The solid curves are theoretical values, the points, experimental ones.

Calculation of the coefficient of reflection

In order to calculate the coefficient of reflection as defined in this paper we must find the area under the P curve (rocking curve) with the radian as the unit of angular measure. Thus

$$R = \Delta\theta_\sigma \frac{\int_{-\infty}^{\infty} [\Phi(\xi_{B\sigma}) + \Phi(\xi_{B\sigma} \sec 2\theta) \cos 2\theta] d\xi_{B\sigma}}{(1 + \cos 2\theta) \int_{-\infty}^{\infty} F(\xi_\sigma) d\xi_\sigma} \tag{32}$$

[17] S. K. Allison, Phys. Rev. **38,** 203 (1931), Eq. (21) 208.

or, finally,

$$R = \frac{\int_{-\infty}^{\infty} \Phi(\xi_{B\sigma}) d\xi_{B\sigma}}{\int_{-\infty}^{\infty} F(\xi\sigma) d\xi_{\sigma}} \frac{1 + \cos^2 2\theta}{1 + \cos 2\theta} 4\delta FZ^{-1} \operatorname{cosec} 2\theta. \tag{33}$$

The values of the integral in the numerator of Eq. (33) have been obtained by graphical integration of curves plotted from the data of Table VI, and are shown in columns 3 and 4 of Table IX. Columns 6, 7, and 8 of Table IX show the calculated and observed coefficients of reflection.

TABLE IX. *Calculation of R from Eq. (33) and comparison with observations.*

Line	λ	$\int_{-\infty}^{\infty} \Phi(\xi_{B\sigma}) d\xi_{B\sigma}$		$\frac{1+\cos^2 2\theta}{1+\cos 2\theta}$	$R \times 10^5$		
		Darwin	Prins		Darwin	Prins	Obs.
W $K\alpha_1$	0.209A	1.631	—	0.999	0.536	—	1.17
Mo $K\alpha_1$	0.708	"	1.537	0.987	1.85	1.84	1.86
Cu $K\alpha_1$	1.54	"	1.238	0.940	3.87	3.49	3.46
Cr $K\alpha_1$	2.28	"	0.917	0.881	5.65	4.53	4.68

A graph of the information in Table IX, together with R values measured by other observers is shown in Fig. 6. Only the experimental data obtained in these experiments on crystals III are included in Fig. 6.

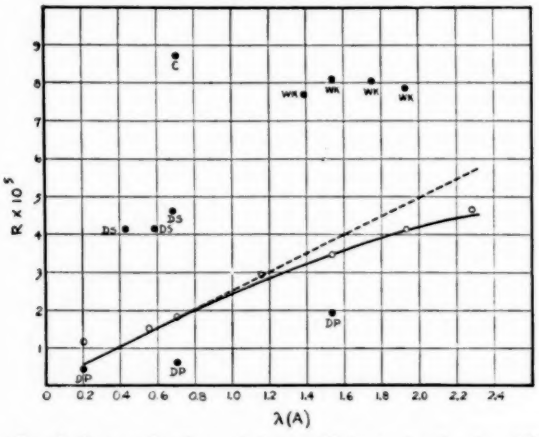

Fig. 6. Calculated and observed values of the coefficient of reflection (R). The dotted curve represents values calculated from Darwin; the solid curve, values from Prins. The open circles are experimental values of crystals III. The solid circles are points by other observers. *C*—Compton;[1] *WK*—Wagner and Kuhlenkampff;[1] *DS*—Davis and Stempel;[1] *DP*—Davis and Purks.[1]

Discussion of results

The most obvious comment to be made on the agreement of observed and calculated coefficients of reflection from 0.5 to 2.3 A as illustrated for crystals III on Fig. 6 is that the agreement is very surprising since no temperature corrections have been applied. The coefficient of reflection should decrease

with rising temperature. The calculation of the temperature correction for a complicated crystal such as calcite is difficult and has not been attempted. The value of $(\sin\theta)/\lambda$ for all the present measurements is 0.165 which is comparatively small, and the temperature correction is proportional to this. There is, however, some evidence that the value $FZ^{-1}=0.47$ used here is too small. This value was based on the structure factors of Bragg and West.[12] More recently James and Brindley[18] and Pauling and Sherman[19] have published tables of structure factors. The exact value of FZ^{-1} depends on the particular electronic structure assumed for the molecule $CaCO_3$ as it appears in the crystal. By assuming arrangements from Ca^{+2}, C^{+4}, O^{-2} to Ca^0, C^0, O^0 we get values of FZ^{-1} which in the table of James and Brindley range from 0.472 to 0.500, and from 0.496 to 0.518 in the table of Pauling and Sherman. If these newer structure factors are preferable to the older values, the value of 0.47 as used in this paper is somewhat too low.[20] It is quite possible that this choice of a low structure factor is compensated by failure to make a temperature correction, thus giving the agreement of Fig. 6. This agreement does not extend to wave-lengths as short as W $K\alpha_1$ (0.2086A), and large deviations are apparent in the quantities plotted in Fig. 5 at this wave-length also. The laws of scattering of these relatively hard wave-lengths are undoubtedly different from the laws applicable in the range 0.5 to 2.3A, a large fraction of the scattering taking place by the Compton process. It is the opinion of the author that this lack of agreement shows that the theory of diffraction by a perfect crystal should be re-investigated in this shorter wave-length region.

If the coefficients of reflection observed by Davis and Purks and entered in Fig. 6 are correct,[21] they present a very serious difficulty, since they are far below the perfect crystal values. The Davis and Purks rocking curve widths for W $K\alpha_1$ $(1, -1)$ were 2.5 seconds full widths at half maximum. The author has been unable to find any calcite specimens that even approach this small value, the values obtained ranging from 4.6 to 14.4 seconds in various specimens. The high values of the coefficient of reflection observed by other authors and entered in Fig. 6 are probably due to the use of calcite surfaces considerably less perfect than those used in the present experiments.

The general statement that these results confirm Prins' modification of Darwin's theory is incorrect in a very important aspect. The present results show that some sort of absorption correction to Darwin's results is needed, but really give no information at all as to the exact shape of the diffraction pattern from a single crystal. In particular, the present results give no information for or against the asymmetrical nature of the patterns in Prins' theory. The fact that the calculated and observed values of P_0 are appreciably dif-

[18] James and Brindley, Phil. Mag. **12**, 81 (1931).

[19] Pauling and Sherman, Zeits. f. Krist. **81**, 1 (1932).

[20] I am indebted to W. H. Zachariasen of this laboratory for stimulating discussions of these matters.

[21] Recent communications from Professor Davis state that on repeating the experiments of Davis and Purks under discussion, the rocking curves from the same sample of calcite were found to be wider than at first reported.

ferent for wave-lengths other than Cr $K\alpha_1$ (at Mo $K\alpha_1$ calculated, 0.776; observed, 0.63) may be taken as evidence that the shape of the single crystal diffraction pattern is not exactly that of Prins. The value of P_0 is quite sensitive to the shape of the pattern, and independent of the polarization. For instance, let us compare three functions all having maxima equal to unity. If the shape of the diffraction pattern for a single crystal were given by $\exp(-\xi^2)$, we should find $P_0=0.71$ from Eq. (26). In the same way if $F(\xi)=(1+\xi^2)^{-1}$, $P_0=0.500$, and if $F(\xi)=(1+\xi^2)^{-2}$, $P_0=0.625$.

A further question of interest is related to the possibility of finding other calcite surfaces as good as those of calcites III in this research. A. H. Compton[3] has reported rocking curves for Mo $K\alpha_1$ (1, −1) of 5 seconds full width at half maximum which agrees with the present results. Hoyt[22] has reported rocking curves under the same conditions somewhat less than 5 seconds wide. Allison and Williams,[1] found a full width at half maximum of 6 seconds. In view of these results it does not seem that such calcite surfaces are exceedingly rare. There is no evidence from the experiments reported here that the width of the rocking curve is a constant for a given piece of calcite, that is, a large sample from a known source. In the author's experiments different splits from the same large piece showed different widths.

Darwin, and other workers following him, have interpreted high values of R, for instance, those of Wagner and Kuhlenkampff in Fig. 6, as evidence for a mosaic structure of crystals. The present research shows that it is possible to obtain calcite surfaces for which there is no evidence of mosaic structure from measurements of the coefficient of reflection by double spectrometer methods in the wave-length region from 0.5 to 2.3A. Presumably the effects of a secondary structure as postulated by Zwicky[23] would not influence the present results.

Note added in proof: Since completion of this article more accurate calculations of theoretical curves have been carried out by eliminating the assumption of Eq. (17). This assumption states that the absorption correction can be rightly applied by assuming that the absorption of an atom in the unit cell is proportional to the fraction of the total number of electrons in the unit, all contained in the atom in question. This is rather far from correct in the absorption by the calcium atoms in calcite for wave-lengths shorter than the K-absorption limit of calcium. The more rigorous calculations direct from Eq. (15) show slight changes in the theoretical values, but do not invalidate in any way the general conclusions of this article. These new calculations will appear in a paper to be submitted from this laboratory in the near future.

[22] A. Hoyt, Paper delivered at the Washington Meeting of the American Physical Society, April, 1932.

[23] Zwicky, Phys. Rev. **40**, 63 (1932).

The Scattering of X-rays by Gases and Crystals

By Y. H. Woo
Department of Physics, National Tsing Hua University, Peiping, China

(Received May 6, 1932)

It is shown that, in exact agreement with the wave mechanical result worked out by Wentzel and Waller and Hartree, the Raman-Compton-Jauncey formula for the scattering of x-rays by an atom deduced on the basis of classical electrodynamics may be written

$$I_\phi = I_e\left\{F^2 + Z - \sum_{r=1}^{z} E_r^2\right\}, \qquad (A)$$

where I_ϕ is the intensity scattered in a direction ϕ with the primary beam, I_e is the scattering from a single isolated electron as calculated by Thomson, Z is the atomic number, E_r is the average amplitude of the waves scattered by the r-th electron in the atom and F is the true atomic structure factor as given by $F=\Sigma_1^z E_r$. This shows that the coherent scattering is proportional to F^2 and that the introduction of the average atomic structure factor F' by Jauncey is unnecessary. On the basis of Eq. (*A*), theoretical formulas for the scattering of x-rays by gases and crystals are redeveloped and the results turn out to explain all the disagreement between theories proposed by Jauncey and the writer.

IN DISCUSSING the scattering of x-rays by polyatomic gases, Jauncey[1] has recently pointed out that the theoretical formula deduced by the writer[2] is not exactly in agreement with that obtained by himself. Since the writer regards the atoms in a polyatomic system (a molecule or a crystal) as the scattering units, this disagreement should also occur when the scattering of x-rays by crystals is dealt with. The purpose of the present note is to discuss this matter in detail.

Taking account of the modification introduced by Jauncey,[3] the Raman-Compton[4] formula for the scattering of x-rays by an atom may be written

$$I_\phi = I_e\{F'^2 + Z - F'^2/Z\}, \qquad (1)$$

where I_ϕ is the intensity scattered at an angle ϕ to a distance R, I_e is the scattering from a single isolated electron as calculated by J. J. Thomson, Z is the atomic number and F' is the average atomic structure factor. According to Jauncey,[3] the relation between the average atomic structure factor F' and the true atomic structure factor F is expressed by

$$F'^2 = F^2 - \left(Z\sum_{r=1}^{Z} F_r^2 - F^2\right)\Big/(Z-1), \qquad (2)$$

[1] G. E. M. Jauncey, Phys. Rev. **39**, 561 (1932).
[2] Y. H. Woo, Phys. Rev. **39**, 555 (1932).
[3] G. E. M. Jauncey, Phys. Rev. **37**, 1193 (1931).
[4] C. V. Raman, Indian J. Phys. **3**, 357 (1928). A. H. Compton, Phys. Rev. **35**, 925 (1930).

where E_r is the average amplitude of the waves scattered by the r-th electron in the atom and where the true atomic structure factor F is given by

$$F = \sum_{r=1}^{Z} E_r. \tag{3}$$

In view of the similarity between Eq. (1) and the formula originally derived by Raman and Compton,[4] the part $I_e F'^2$ was taken[5] by analogy to represent the coherent scattering and the other terms to represent the incoherent scattering. However, there exists no physical or mathematical reason for justifying this separation. On substituting the value of F' given by (2) into Eq. (1), we find as a result of straightforward calculation

$$I_\phi = I_e \left\{ F^2 + Z - \sum_{r=1}^{Z} E_r^2 \right\}. \tag{4}$$

This is exactly the same expression as that obtained by Wentzel[6] and independently by Waller and Hartree[7] on the basis of wave mechanics. This shows that the coherent part of the total intensity is proportional to F^2 and that the introduction of the average atomic structure factor F' by Jauncey is unnecessary. When corrected for the change of wave-length, Eq. (4) becomes

$$I_\phi = I_e F^2 + I_e \frac{Z - \sum_{1}^{Z} E_r^2}{[1 + \gamma(1 - \cos\phi)]^3}, \tag{5}$$

where $\gamma = h/mc\lambda$ and h, m, c and λ have their usual significance.

Recently Herzog[8] has actually compared Eq. (5) with experiments reported by Wollan[9] on absolute measurements of the scattering of Mo K_α rays by helium, neon and argon gases. The agreement seems to be quite satisfactory.

On the basis of Eq. (5) and following the arguments presented by the writer in previous papers,[10] the theoretical formulas for the scattering of x-rays by gases and crystals can be easily redeveloped. For the scattering from a simple cubic crystal consisting of atoms of one kind, the diffusely scattered intensity is given by

[5] Cf. G. E. M. Jauncey, Phys. Rev. **38,** 1 (1931); **39,** 561 (1932); Y. H. Woo, Phys. Rev. **39,** 555 (1932).

[6] G. Wentzel, Zeits. f. Physik **43,** 1 and 779 (1927).

[7] I. Waller, Zeits. f. Physik **51,** 213 (1928); Waller and Hartree, Proc. Roy. Soc. **A124,** 119 (1929). The more exact formula obtained by these authors differs from Eq. (4) by including some negative terms arising from consideration of Pauli's principle, but the contribution to the total scattering due to these negative terms is negligible in comparison with that due to the terms included in Eq. (4).

[8] G. Herzog, Zeits. f. Physik **70,** 583 and 590 (1931). Cf. also Y. H. Woo, Sci. Rep. Tsing, Univ. AI, No. 4 (1932).

[9] E. O. Wollan, Phys, Rev. **37,** 862 (1931).

[10] Y. H. Woo, Proc. Nat. Acad. Sci. **17,** 467 (1931); Phys. Rev. **38,** 6 (1931), and **39,** 555 (1932).

$$I_\phi = NI_e\{(1 - e^{-2M})F^2 + \frac{Z - \sum_1^Z E_r^2}{[1 + \gamma(1 - \cos\phi)]^3}. \quad (6)$$

For the intensity scattered by a gas molecule containing atoms $I, \cdots i, \cdots n$ arranged at fixed distances from each other, the formula becomes

$$I_{\phi n} = I_e \sum_1^n \sum_1^n F_i F_j \frac{\sin kS_{ij}}{kS_{ij}} + I_e \sum_1^n \frac{Z_i = \sum_1^{Z_i} E_r^2}{[1 + \gamma(1 - \cos\phi)]^3}. \quad (7)$$

In Eq. (6) N is the number of atoms per unit volume of the crystal and e^{-2M} is the temperature factor as calculated by Debye[11] and Waller[12] and in Eq. (7) F_i is the true atomic structure factor for the i-th atom, $k = (4\pi/\lambda)\sin\frac{1}{2}\phi$, s_{ij} is the distance of any atom i from any other atom j. The other quantities have the same meaning as those employed in Eq. (5). When applied to the scattering by a diatomic gas consisting of two like atoms, Eq. (7) gives for the scattering per electron

$$S = \frac{I_{\phi 2}}{2ZI_e} = \left(1 + \frac{\sin kS}{kS}\right)\frac{F^2}{Z} + \frac{1 - \frac{1}{Z}\sum_1^Z E_r^2}{[1 + \gamma(1 - \cos\phi)]^3} \quad (8)$$

where s is the distance between the two atoms in the scattering molecule.

It will be noticed that Eq. (6) agrees with the formula deduced by Jauncey and Harvey[13] provided the latter is properly corrected for the change of wavelength. Also a comparison of Eqs. (7) and (8) with those recently obtained by the writer[14] and Jauncey[15] will clearly show that, if the coherent and incoherent scattering be properly separated, the disagreement pointed out by Jauncey no longer exists. Finally, referring the comparison between theory and experiment recently made by the writer,[14] it can be readily concluded that Eq. (8) is supported by Wollan's measurements[16] on the scattering of Mo K_α radiation by diatomic gases.

[11] P. Debye, Ann. d. Physik **43,** 49 (1914).
[12] I. Waller, Zeits. f. Physik **17,** 389 (1923); *Upsala Dissertation,* (1925).
[13] Jauncey and Harvey, Phys. Rev. **37,** 1203 (1931).
[14] Y. H. Woo, Phys. Rev. **39,** 555 (1932).
[15] G. E. M. Jauncey, Phys. Rev. **39,** 561 (1932).
[16] E. O. Wollan, Phys. Rev. **37,** 862 (1931); Proc. Nat. Acad. Sci. **17,** 475 (1931).

Ionization of Air by γ-Rays as a Function of Pressure and Collecting Field

By I. S. Bowen
Norman Bridge Laboratory of Physics, California Institute of Technology

(Received May 23, 1932)

The ionization of air by γ-rays was studied for pressures from 1 to 93 atmospheres and for collecting fields from 1.55 to 1009 volts per cm. Increases in ionization current of over 40 percent were observed when the potential gradient was varied over this range, thus indicating that the lack of proportionality of ionization current with pressure obtained by previous observers was principally due to lack of saturation.

I. Introduction

THE ionization of air by γ-rays or by cosmic rays has been studied through a wide range of air pressures by Downey,[1] Fruth,[2] Broxton,[3] Swann,[4] and Millikan.[5] All of these observers agree in finding that as high pressures are reached the ionization per atmosphere decreases to a small fraction of its value at one atmosphere. In fact, the total ionization does not increase appreciably as the pressure is pushed above 100 atmospheres.

An explanation of this decrease has been proposed by Downey and elaborated by Broxton as follows. Practically all of the ionization in a gas is caused, not by the γ-rays or the cosmic rays themselves, but by the secondary β-rays ejected by them. At atmospheric pressure in any vessel of ordinary size the β-rays traversing the chamber are formed largely in the walls of the vessel. As the pressure is increased, however, the ranges of these β-particles decrease until they finally become less than the dimensions of the chamber. When this condition is reached the β-rays from the walls have produced all of the ions that they are capable of, and therefore any further increase of pressure cannot increase the total ionization caused by them. This theory appears then to predict the ionization pressure curves obtained by all observers.

However, it is a well established fact that the mass absorption coefficients of γ-rays and cosmic rays are not widely different for light substances such as air or water from their values for the metals out of which ionization chambers are usually made. Since the γ-rays or cosmic rays are absorbed in the air and since the only way in which their energy can be absorbed is by transfering it to secondary β-rays, it is obvious that the air must have approximately the same efficiency of production of β-rays as the walls. This being the case, any β-rays from the wall that are absorbed by the air are replaced by a corresponding number of β-rays ejected from the air itself. If

[1] K. M. Downey, Phys. Rev. **16,** 420 (1920); **20,** 186 (1922).
[2] H. F. Fruth, Phys. Rev. **22,** 109 (1923).
[3] J. W. Broxton, Phys. Rev. **27,** 542 (1926); **28,** 1071 (1926); **37,** 1320 (1931).
[4] W. F. G. Swann, J. Frank. Inst. **209,** 151 (1930).
[5] R. A. Millikan, Phys. Rev. **39,** 397 (1932).

this is true, the number of ions formed should be directly proportional to the pressure.

In order to reconcile this conclusion with the failure of all observers to collect this large a number of ions at the higher pressures, one is at once led to the suggestion, made independently by Millikan and Bowen[6] and by Compton, Bennett and Sterns,[7] that saturation was not attained at these high pressures, in spite of the fact that all experimenters tested for saturation in the usual way.

In all of these experiments the ionization chamber was a cylinder or sphere with a small collecting rod or fibre in the center. With this arrangement most of the potential drop is near the collecting rod while the great bulk of the chamber is subject to a potential gradient that, expressed in volts per cm, numerically amounts to only a very few percent of the total collecting potential applied. For this reason the highest potentials that these observers found it feasible to use, gave a gradient only a little above the point at which the ionization current became even approximately constant as the gradient was varied. Thus Broxton, in testing for saturation, varied his potential gradient through only 20 percent while the other experimenters carried their tests to only two or three times the gradient where the current became approximately constant. Obviously these tests would not bring to light variations of the type found in the experiments described in this paper.

II. Description of Apparatus

As no indication of a lack of saturation was found by previous observers when the collecting potential was varied through a small range, it was necessary to build an ionization chamber in which very much higher potential gradients could be applied if a satisfactory test of this lack of saturation was to be obtained. Obviously the older type of chamber was unsuited to this because of the very high total potentials necessary. Consequently a chamber was constructed in which the ions were collected between parallel plates, thus making possible a high uniform field. The ionization chamber was a steel cylinder 12.5 cm in diameter and 20 cm in length, designed to hold a pressure of 100 atmospheres. In this were placed 8 plates 10 cm in diameter which could be connected to a source of high potential. Alternating with these were 7 collecting plates 7.5 cm in diameter. These collecting plates were surrounded by grounded guard rings from which they were insulated by small amber lugs in grounded metal shields. The inside diameter of the guard rings were 7.7 cm and the outside 10 cm. The collecting plates were separated from the potential plates by 1 cm. This arrangement made possible the collection of ions from a definite known volume in which there was a uniform potential gradient which was numerically equal to the potential applied to the potential plates.

Ideally it should have been possible to measure directly the ionization by

[6] R. A. Millikan and I. S. Bowen, Nature **128,** 582 (Oct. 3, 1931), See also reference 5.

[7] A. H. Compton, R. D. Bennett and J. C. Sterns, Phys. Rev. **38,** 1565 (Oct. 15, 1931) and Phys. Rev. **39,** 873 (1932).

connecting an electrometer to the collecting plates and observing their rate of change of potential with time. Unfortunately however a change in potential of the high potential plates due to the unavoidable variation in the battery used to maintain it, produces by induction a change in potential of the collecting plates. Thus because of the large electrical capacity of the system, the ionization at the lowest pressure used, caused a change of only 0.3 volt per hour in the potential of the collecting plate. It was obviously impossible to maintain a 1000 volt battery so constant that its variation would not be an appreciable fraction of this. To compensate for this variation a second set of plates was built. This was an exact duplicate of the first set except that it was enclosed in a light brass case. The high potential plates of the ionization chamber A and the compensating chamber B were then connected to the same battery as shown in Fig. 1, while the electrometer E was arranged to measure

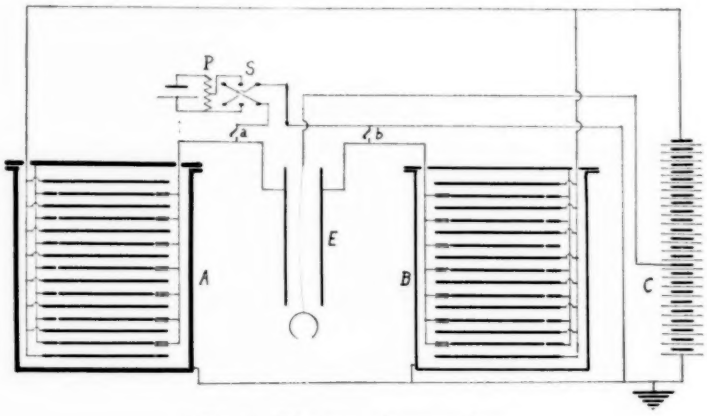

Fig. 1. Diagram of connections.

the difference in potential of the two sets of collecting plates. The electrometer used was of the single string type. The two plates of the electrometer were connected respectively to the two sets of collecting plates, while a potential of 225 volts was applied to the string. The electrometer was adjusted for a sensitivity of about 500 divisions per volt.

Because of the fact that cosmic rays at sea level cause less than $\frac{1}{4}$ of the ionization in a closed vessel, the remainder being due to γ-rays from radioactive materials in the neighborhood, and to α-rays from radioactive impurities in the walls of the ionization chamber, it was clear that more interpretable results could be obtained if the ionization of a single source such as γ-rays were studied.

For this purpose γ-rays from a small quantity of radio-thorium, filtered through 13 mm of lead were used. In order to obtain a well defined beam of γ-rays the radio-thorium and filter were placed in the center of a lead cylinder 30 cm in diameter and 30 cm long. There was a conical opening leading to one end of the cylinder which allowed a beam to pass that was somewhat larger than the ionization chamber at its position about 2 meters away. The block of lead was on a swivel so that the beam of γ-rays could be directed

either at the ionization chamber or turned to one side of it. By taking readings with the γ-rays passing through the chamber and then again with it turned to one side it was possible to determine the ionization of the γ-rays independent of the ionization of cosmic rays or α-rays.

The usual procedure in taking readings was to connect the collecting plates of the compensating chamber B to ground while the collecting plates of the ionization chamber A were held at some potential, usually 0.05 volts below ground, by means of the potentiometer P. The electrometer was then read. Next the connections a and b were broken, thus leaving both sets of collecting plates floating. The electrometer was then immediately read again, and the stop watch simultaneously started. After approximately enough time had elapsed for the collecting plates of the ionization chamber to receive enough ions to bring their potential to the same amount above that of the collecting plates of the compensating cylinder as they had started below, the electrometer position was again noted and the stop watch read. Switches a and b were now closed and this same potential above ground was then applied to the plates by means of the reversing switch S and the electrometer again read. In this way it was possible to correct for any zero drift or change in sensitivity of the electrometer.

This gave then with a considerable accuracy the rate of change of potential difference between the collecting plates. The relationship between this rate of change of potential and the ionization per cc per sec. was obtained as follows: Let C_1 be the electrical capacity between the collecting plates and the high potential plates of the ionization chamber, C_2 the capacity to ground of the collecting plates, leads and electrometer plates. By symmetry these capacities are the same for compensating chamber. Let C_3 be the capacity between the plates of the electrometer. It is then easy to show that if a charge Q is accumulated on the collecting plates of the ionization chamber and Q_2 on these plates of the compensating chamber the potential difference in volts across the electrometer is W where

$$W = (300(Q_1 - Q_2))/(C_1 + C_2 + 2C_3). \tag{1}$$

Likewise it may be shown that if the potential plates of the compensating cylinder are grounded while the potential of these plates in the ionization chamber are varied by an amount v, then the charges induced on the collecting plates (left floating) are such as to change the potential between the collecting plates of the two cylinders by an amount w, where

$$w = vC_1/(C_1 + C_2 + 2C_3). \tag{2}$$

w is obviously the change in reading of the electrometer when this variation takes place and hence can be determined directly.

From (1) it is seen that the difference of the ionization currents in the two chambers is

$$\frac{dQ_1}{dt} - \frac{dQ_2}{dt} = \frac{dW}{dt} \cdot \frac{C_1 + C_2 + 2C_3}{300} \tag{3}$$

where dW/dt is the rate of change of potential across the electrometer, in volts per sec. But substituting from (2)

$$\frac{dQ_1}{dt} - \frac{dQ_2}{dt} = \frac{dW}{dt} \cdot \frac{vC_1}{300w}. \tag{4}$$

Since, however, the distance between the potential and collecting plates is 1 cm

$$C_1 = A/4\pi$$

where A is the total effective area of the collecting plates in each cylinder

$$\frac{dQ_1}{dt} - \frac{dQ_2}{dt} = \frac{dW}{dt} \cdot \frac{v}{300w} \cdot \frac{A}{4\pi}.$$

The difference in the number of ions $(N_1 - N_2)$ collected per cc per sec. in the two chambers is then, since the volume from which ions are obtained is A cc and the charge on each ion is e,

$$N_1 - N_2 = \frac{1}{eA}\left(\frac{dQ_1}{dt} - \frac{dQ_2}{dt}\right) = \frac{dW}{dt} \cdot \frac{v}{w300 \times 4\pi e}.$$

By taking the difference of the readings with and without the γ-rays, the N_2 cancels out and the number produced by the γ-rays alone are obtained.

Since in these experiments γ-rays giving an ionization current about 20 times as large as the natural ionization current caused by cosmic rays, α-rays etc. were used, it was advisable to test whether the form of the ionization-voltage curves was dependent on the intensity of the radiation. For this purpose determinations of the ionization currents were made, at the two highest pressures, when the γ-rays were cut down to less than $\frac{1}{5}$ of their normal intensity by being passed through 5 cm of steel.

The pressures up to 25 atmospheres were measured on a pressure gauge that was calibrated, during the course of the experiments, against a gauge tester. Consequently these values should be in error by less than 0.1 atmosphere. It was not feasible to calibrate the gauge used on the highest pressure and therefore there may be a rather large error in this pressure.

In all cases a given collecting potential was maintained on the collecting plates for several hours before readings were started to allow any charges accumulated on the insulators to take up a steady value.

III. Results

The ionization currents observed in these experiments for a series of pressures and potential gradients are given in Table I. The ionization current is expressed as the number of ions per cc per sec. per atmosphere. For the two highest pressures two columns of values are given. The first column marked I_1 represents the ionization at the full intensity of the γ-rays, that is, the same intensity as that used for the lower pressures. The second column marked 5.187 I_2 is 5.187 times the number of ions obtained when the intensity of the γ-rays was reduced by passage through 5 cm of iron. Except for the two lowest gradients at the 93 atmospheres pressure, the two columns agree within the experimental error, thus indicating that with these two exceptions

TABLE I. *γ-rays. Number of ions collected per cc per second per atmosphere.*

Field in volts per cm \ Pressure in atm.	0.98	3.74	10.50	24.95 I_1	24.95 5.187 I_2	93. I_1	93. 5.187 I_2
1.55	112.5	101.6	79.8	59.8	59.8	26.2	28.4
6.2	118.4	106.9	86.2	63.9	63.7	29.7	30.3
23.0	118.7	110.1	89.5	66.5	66.5	30.6	30.5
91.5	120.0	113.9	94.5	70.2	70.4	32.1	31.9
367.	119.7	116.8	102.1	77.1	77.3	34.8	35.0
1009.	121.1	118.1	108.4	85.5	85.8	39.1	39.1

the form of the curves is independent of the intensity for intensities up to those used.

Fig. 2 displays graphically these data, the number of ions per cc per sec. per atmosphere being plotted against the log of the potential gradient of the collecting field P. It is at once seen that, in agreement with previous observers, the ionization per atmosphere falls off rapidly at the high pressures. Furthermore it agrees with their observation that when the potential gradient is of the order of 5 or 10 volts per cm, doubling this gradient produces a very small change in the current.

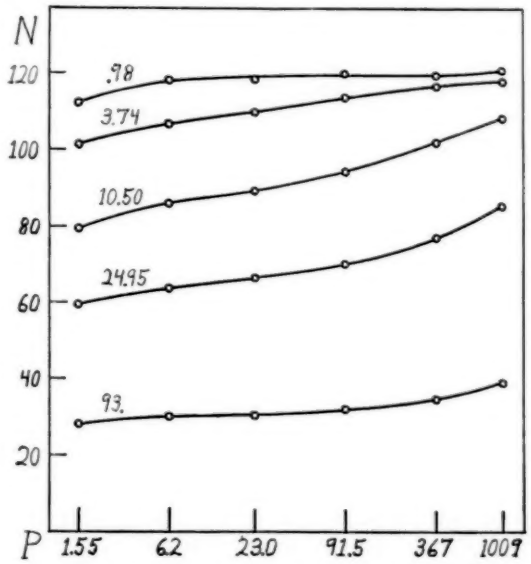

Fig. 2. The number of ions per cc per sec. per atmosphere (N) collected as a function of the potential gradient in volts per cm P for the pressures in atmospheres indicated.

However when high potential gradients are applied a very material increase of the currents is obtained thus showing that the lower gradients do not collect all of the ions formed. Thus if we assume that the number of ions per atmosphere actually formed is constant and has the value 121.1 ions per cc per sec., then the number that one fails to collect at a gradient of 6.2 volts per cm (approximately the field used by previous observers) is $(121.1 - N_{6.2})$. The fraction of these that one still fails to collect at some other higher poten-

tial P is $R=(121.1-N_P)/(121.1-N_{6.2})$. This fraction then may be taken as a measure of the success attained in pushing towards saturation by using higher fields. This fraction R as calculated from Table I is plotted against the log of P in Fig. 3.

From the figure it is seen that at 3.74 atmospheres it is possible by increasing the gradient from 6.2 to 1009 volts per cm to collect all but 21 percent of the ions that are not obtained with the lower gradient. At the higher pressures the increase of the field strength is progressively less effective.

These results provide conclusive evidence that the falling off of ionization per atmosphere at higher pressures is largely due to lack of saturation rather than to the mechanism suggested by Broxton.[8]

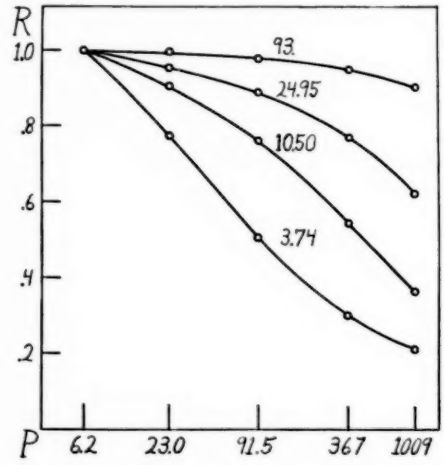

Fig. 3. $R=(121.1-N_P)/(121.1-N_{6.2})$ as a function of the potential gradient P for the pressures indicated.

Furthermore the lack of dependence of the form of the curves on the intensity of ionization, in nearly all cases, indicates that the lack of saturation is caused by recombination of the ejected electron with the parent ion rather than by random recombination between any ions formed. Presumably at the high pressures the ejected electron loses its energy before it succeeds in escaping from the region where the field of the parent ion is still very strong. This fortunately also means that this lack of saturation does not invalidate the results of many cosmic ray observers who have used high pressures, since the lack of saturation merely cuts down all currents by a constant factor and in no way effects the relative values on which absorption coefficients are based.

This increase of ionization current with potential gradient also seems to call for a modification of the theory developed by Compton, Bennett and Sterns[7] to explain this lack of saturation. According to their theory there is a

[8] There may still be a slight variation of the type suggested by Broxton due to a small difference in efficiency of production of β-rays by the air and by the steel walls of the chamber. Unpublished studies on the effect of wall materials by Workman in this laboratory indicate however that this effect of these materials is at most 30 percent and is probably in the opposite direction, i.e., it should cause the ionization per atmosphere to increase slightly with pressure.

certain critical sphere of definite radius surrounding each ion. If an oppositely charged ion finds itself inside of this sphere the force of attraction over balances the tendency to diffuse away and the two ions always recombine. On the other hand the tendency to diffuse away predominates for an ion outside of this sphere and the ion always escapes. Since at the surface of this critical sphere the field due to the ion at the center is 40,000 volts per cm any ordinary collecting field should have no effect. The question whether recombination with the parent ion takes place or not is determined, according to this theory, solely by the distance from the parent ion that the electron is ejected by the β-particle, and the factors that enter into the tendency to diffuse away such as the temperature.

Obviously this picture is much too simplified. Thus given two ions in a gas there is always a certain probability that one will diffuse close enough to the other, regardless of their original distance apart, so that recombination takes place. Of course this probability falls off rapidly as the distance increases. Furthermore this probability, particularly for an ion at a considerable distance from the parent ion, is materially changed as the strength of an external collecting field is varied. This of course changes the number that succeed in escaping from the parent ion and therefore one would expect variation in the strength of the ionization current similar to that shown in the Table I and Fig. 2.

Qualitatively this is also in agreement with the behavior indicated in Fig. 3. At the lower pressures all of the ions start at a great distance from the parent ion where its field is small and consequently the probability of recombination should be largely effected by the size of the external field, as it is found to be. On the other hand for higher pressures most of the ions start so close to the parent ion where its field is large compared to that of any feasible external field. This being the case very little effect of the strength of the external field should be expected.

TABLE II. *Residual ionization. Number of ions collected per cc per second per atmosphere.*

Field in volts per cm \ Pressure in atm.	24.95	93.
1.55	3.11	1.34
6.2	3.22	1.45
23.0	3.38	1.55
91.5	3.53	1.64
367.	4.00	1.77
1009.	4.42	1.99

The readings made when the γ-rays were turned to one side represent the difference in the residual ionization caused in the two chambers by cosmic rays and radioactivity of the chamber and room. This was large enough in the case of the upper two pressures to be read with a considerable accuracy. Also in the case of these high pressures, the ionization in the ionization chamber was large compared to that in the compensating chamber and consequently the fact that the latter is subtracted should not modify the values very greatly. These results are given in Table II. As is at once seen the effect of increase in collecting potential is essentially the same as with the γ-rays.

A Search for Isotopes of Hydrogen and Helium

By WALKER BLEAKNEY*
Palmer Physical Laboratory, Princeton University

(Received June 1, 1932)

The well-known triatomic hydrogen ion $(H^1H^1H^1)^+$ of mass 3 is formed by electron impact at a rate proportional to the square of the pressure if the pressure is very low. The isotopic molecular ion $(H^1H^2)^+$, also of mass 3, would be expected to vary linearly with pressure. These two facts constitute a method for distinguishing between the two with a mass spectrograph whose resolving power is insufficient to separate them by magnetic analysis. The method becomes successful only when the working pressures are very low. With a mass spectrograph designed for low pressure work evidence was found which amply confirms the existence of an isotope of hydrogen of mass 2 and also gives a fair estimate of the abundance. In a sample of ordinary commercial electrolytic hydrogen the data indicate an abundance ratio

$$H^2/H^1 = 1/30000 \pm 20\%.$$

The same ratio for a sample of hydrogen which had been concentrated by Urey, Brickwedde and Murphy came out to be

$$H^2/H^1 = 1/1050 \pm 5\%.$$

An effort was made to detect a difference in mass between the ions $(H^1H^1H^1)^+$ and $(H^1H^2)^+$ without success. Considering the resolving power of the apparatus this failure was interpreted as meaning that the packing fraction of H^2 is greater than 4×10^{-3} i.e., the atomic weight is greater than 2.008. A search was also made for isotopes 3 and 5 of helium but none were found. It is concluded that their abundance must be less than one part in 50,000.

INTRODUCTION

THE hydrogen and helium atoms play such fundamental roles in all physics that any new properties or characteristics of their nuclei are of wide interest. There have been indications that two or more isotopes of these elements may exist and several attempts have been made from time to time to find them experimentally. The first positive evidence for the existence of an isotope of hydrogen of mass 2 has been obtained recently by Urey, Brickwedde and Murphy[1] whose paper also gives a review of previous work and a bibliography of the subject which need not be repeated here. Their photographs of the Balmer spectrum under very high dispersion reveal weak satellite lines attributable to hydrogen atoms of mass 2. These lines were enhanced in a sample of hydrogen in which enrichment of the heavy isotope had been attempted by evaporation near the triple point. Further evidence for the existence of an isotope in hydrogen may be found in Allison's interpretation of his experiments on optical rotation[2] but his method is as yet little

* National Research Fellow.

[1] H. C. Urey, F. G. Brickwedde and G. M. Murphy, Phys. Rev. **40**, 1 (1932); also Phys. Rev. **39**, 164 (1932) and **39**, 864 (1932).

[2] F. Allison, Jour. Ind. Eng. Chem. **4**, 9 (1932).

understood. The importance of the problem seemed to warrant an independent investigation from a different angle and accordingly a careful reexamination of the hydrogen and helium ions was undertaken with a mass spectrograph.

There are many complications which arise in the magnetic analysis of hydrogen ions because of the various primary and secondary combinations involved. Let us assume that a rare isotope of mass 2 does exist and let us enumerate the possibilities when the ions are formed at a very low pressure by electron impact. It is well known[3] that under these conditions monatomic, diatomic and triatomic ions are produced. Hence we should expect the various combinations listed in Table I. The column labeled $f(p)$ gives the intensity or number of ions as a function of pressure when p is so small that

TABLE I. *Possible ions resulting from a mixture of two hydrogen isotopes in which H^1 is much more abundant than H^2.*

	Ion	m/e	$f(p)$	Intensity
1.	$(H^1)^+$	1	$a_1p+b_1p^2$	weak
2.	$(H^1H^1)^+$	2	a^2p	v. strong
3.	$(H^2)^+$	2	$a_3p+b_3p^2$	v. weak
4.	$(H^1H^1H^1)^+$	3	b_4p^2	weak
5.	$(H^1H^2)^+$	3	a_5p	weak
6.	$(H^2H^2)^+$	4	a_6p	v.v. weak
7.	$(H^1H^1H^2)^+$	4	b_5p^2	v.v. weak
8.	$(H^1H^2H^2)^+$	5	b_6p^2	v.v. weak
9.	$(H^2H^2H^2)^+$	6	b_7p^2	v.v. weak

secondary reactions are small compared with primary ones. In general the primary processes will vary linearly with pressure while secondary ones will be proportional to the square. Of the configurations listed in Table I it is to be expected that 6, 7, 8 and 9 will be too weak for observation. Number 1 is of no interest in this particular case and 3 will be completely masked by 2. Since 2 is a linear function of pressure and since it is very strong it serves as a good index of the pressure. There are left 4 and 5, two ions of comparable abundance but different functions of pressure and it is in this direction that the following studies were made. In the case of helium no such complications arise.

EXPERIMENTAL PROCEDURE

The apparatus used in this experiment has been described[4] in a previous number of this journal and some preliminary results have also been reported.[5,6] The flow method was used in every case, the gases streaming into the tube through fine capillaries or a palladium tube and out through diffusion pumps. The lighter the atom the faster will it leak through the capillary but presumably the same selective effect will occur in the diffusion pumps so that the relative concentrations of isotopes in the ionization chamber will

[3] H. D. Smyth, Rev. Mod. Phys. **3**, 347 (1931).
[4] W. Bleakney, Phys. Rev. **40**, 496 (1932).
[5] W. Bleakney, Phys. Rev. **39**, 536 (1932) (Letter).
[6] W. Bleakney, Phys. Rev. **40**. 130 (1932) (Abstract).

remain unchanged after equilibrium has once been established. No difference in the results on hydrogen were observed when the palladium tube was substituted for the fine capillary.

Through the kindness of Professor Urey and Dr. Murphy a sample of the residue of hydrogen evaporated near the triple point by Dr. Brickwedde at the Bureau of Standards was provided. The details of the preparation of this hydrogen have been described in their paper.[1] This particular sample was designated by them "Sample III." Ordinary commercial electrolytic hydrogen was used for the rest of the work.

The procedure in the case of hydrogen was to measure the total number of ions of mass 3 reaching the collector, per unit electron current, as a function

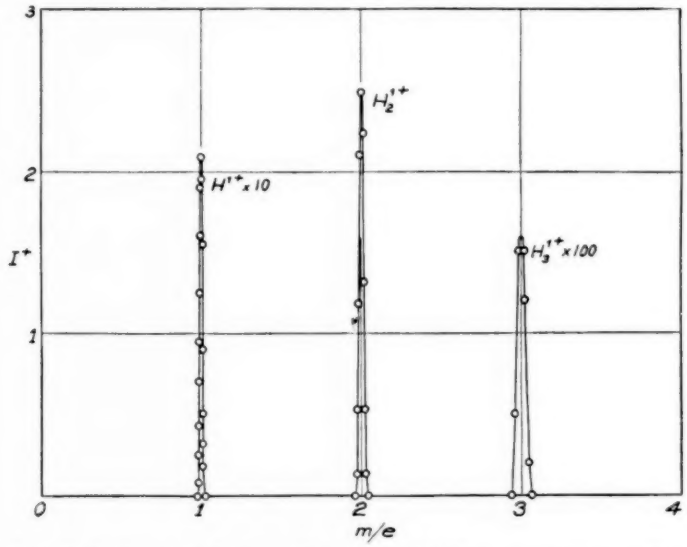

Fig. 1. A typical e/m analysis curve for ordinary hydrogen. p 10^{-5} mm Hg. H = 400 gauss. V_a = 60 volts electron velocity.

of the pressure. The pressure itself could be measured in two ways. The first method consisted in observing the intensity of the ion peak of mass 2 which is proportional to p. In the second method the tube was used as an ionization gauge by observing the total positive ion current per unit electron current. The two methods gave substantially the same results. At the low pressures used the McLeod gauge could not be read with any accuracy. After baking the tube the residual pressure of foreign gases was always in the neighborhood of 10^{-8} mm. The working pressures ranged from 10^{-4} to 10^{-6} mm.

Results for Hydrogen

Some idea of the resolving power of the apparatus may be had from Fig. 1 which shows the three typical peaks obtained in ordinary hydrogen with a magnetic field of 400 gauss and a pressure of about 10^{-5} mm. The electron velocity was 60 volts. At higher magnetic fields the peaks become somewhat sharper. Fig. 2 represents the number I of ions of mass 3 as a function of the number of ions of mass 2. Both ordinate and abscissa are plotted in the same

arbitrary units. Curve I was observed for ordinary hydrogen and curve III for Brickwedde's concentrated sample. The difference between these two curves is given by the straight line II.

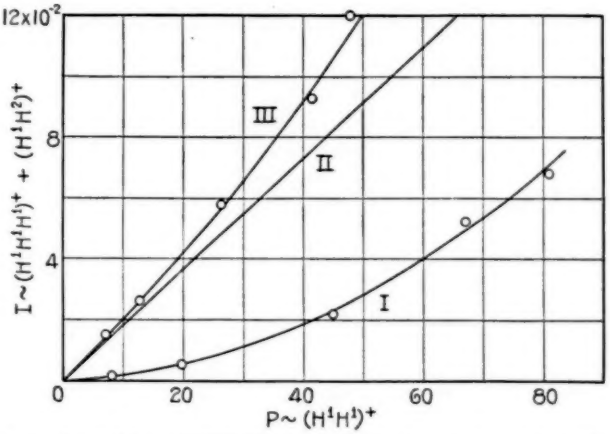

Fig. 2. Total number of ions $(H^1H^1H^1)^+ + (H^1H^2)^+$ of mass 3 as a function of the number $(H^1H^1)^+$ of mass 2. Curve I, ordinary commercial electrolytic hydrogen. Curve III Brickwedde's concentrated sample. Curve II, difference.

These data are interpreted by the author in the following manner. As indicated in Table I we should expect the number of ions of mass 3 to be given by

$$I = ap + bp^2. \tag{1}$$

Now it is significant that curves I and III may be fitted by Eq. (1) using a different value for the constant a but the same value for b in each case. The circles represent the experimental points while the curves are plotted from the following equations.

$$\text{Curve I} \qquad I_{\mathrm{I}} = (\ 6.6p - p^2) \times 10^{-5} \tag{2}$$

$$\text{Curve III} \qquad I_{\mathrm{III}} = (190.0p - p^2) \times 10^{-5} \tag{3}$$

$$\text{Curve II} \quad I_{\mathrm{II}} = I_{\mathrm{III}} - I_{\mathrm{I}} = 183.4 \times 10^{-5}p. \tag{4}$$

The left-hand member of Eq. (4) is interpreted as the difference between the numbers of isotopic ions $(H^1H^2)^+$ in the two samples. Since p is actually measured in terms of the number of $(H^1H^1)^+$ ions it is evident that the slope of curve II gives the increase due to the concentration process.

$$H^1H^2/H^1H^1 = 1/546. \tag{5}$$

Here we assume of course that the probabilities of ionization of the two types are the same. The total concentration of the H^1H^2 molecules is given by the coefficient a of Eq. (1). For curve III this is

$$H^1H^2/H^1H^1 = 1.90 \times 10^{-3} = 1/526 \tag{6}$$

and for ordinary hydrogen

$$H^1H^2/H^1H^1 = 6.6 \times 10^{-5} = 1/15100. \quad (7)$$

The probable error in these two results is estimated to be 5 percent and 20 percent respectively. No attempt was made to treat the data by the method of least squares in detail since neither the ordinate nor the abscissa is known precisely and the probable error of each observation varies from point to point. The curves were therefore fitted in a rather arbitrary way. If we write the ratio of numbers of atoms instead of molecules we have

$$H^1/H^2 = 1050 \pm 5\% \quad (8)$$

for the concentrated sample and

$$H^1/H^2 = 30{,}000 \pm 20\% \quad (9)$$

for ordinary hydrogen. It will be noticed that there is quite a marked discrepancy between these results and those first reported by Urey, Brickwedde and Murphy,[1] but they have largely succeeded in accounting for the difference[7] by considering the relative absorption of strong and weak lines in their discharge tube.

In order to show the results in a more striking manner it is illuminating to divide Eq. (1) by the pressure and so obtain

$$I/p = a + bp \quad (10)$$

an equation of a straight line whose intercept a gives the abundance of the isotope sought. Curves I and III of Fig. 3 plotted in this manner represent the same data as are shown in Fig. 2. This method of presenting the data improves one's confidence in the evidence for the existence of the isotope in ordinary hydrogen but on the same scale the points for the concentrated sample appear scattered over a considerable range. However a little consideration will probably convince the reader that the probable error chosen is not too small. The curves I and III are plotted from Eqs. (2) and (3).

It is of importance to point out that the triatomic ion $(H^1H^1H^1)^+$ is not only a function of the pressure but also of the geometry of the apparatus and the fields applied. The coefficient b of Eq. (1) is related to the intensity of this ion and is given by the slope of the curves in Fig. 3. The relative number of isotopic molecules, however, should be independent of these conditions. Curve IV was taken with an altered set of electric and magnetic fields and it is evident that the slope b has been greatly increased while the intercept remains unchanged within the experimental error. The fact that curves I and IV have exactly the same intercept is not a coincidence. Since their probable errors overlap they were *made* to coincide by the author in fitting the curves.

Some data for another sample of hydrogen are shown by curve V in Fig. 3. This sample was prepared by Professor Urey by a diffusion method and represents a concentration of the lighter isotope. The data are admittedly rather inaccurate but the results seem to indicate less of the heavier isotope

[7] H. C. Urey, F. G. Brickwedde and G. M. Murphy, Phys. Rev. **40**, 464 (1932).

than is to be found in ordinary hydrogen. Professor Urey estimated the concentration of H^1H^2 in this sample to be about one tenth of that in ordinary hydrogen, a result which is not inconsistent with the somewhat higher one obtained here when account is taken of the rather large probable error involved in locating the intercept of curve V.

Kallmann and Lasareff[8] have attempted to verify the existence of the H^2 isotope by seeking for the $(H^1H^1H^2)^+$ ion of mass 4, but were unable to find it.

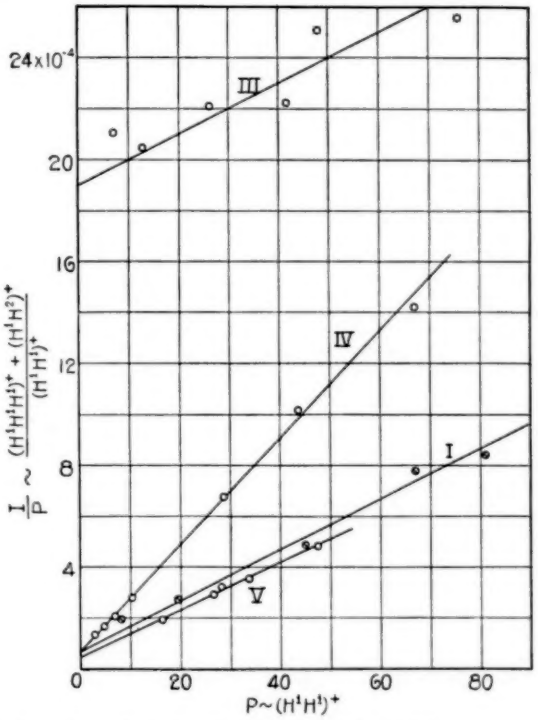

Fig. 3. I/p plotted as a function of p. Curve III, Brickwedde's concentrated sample. Curves I and IV, ordinary commercial electrolytic hydrogen each taken under a different set of fields. Curve V, sample of hydrogen enriched in H^1 by diffusion process.

Perhaps the probability for its formation is very small. They then tried the method of this experiment but at the lowest pressures possible in their apparatus (5×10^{-5}) the ratio of ions of mass 3 to those of mass 2 was 1:4000 for ordinary and concentrated samples alike. The conclusion they reach is that their concentrated sample is at most one tenth as good as the one used in this work. If this is true then there is probably no contradiction between the two experiments.

An accurate determination of the mass of the new isotope is of the greatest importance. Unfortunately the mass spectrograph employed has not a high dispersion on the mass scale. Nevertheless a trial was made to determine if possible a lower limit for the packing fraction. First ordinary hydrogen was admitted and the shape of the peak corresponding to ions of mass 3 was ex-

[8] H. Kallmann and W. Lasareff, Naturwiss. **20,** 206 (1932).

amined very carefully. The pressure was such that this peak was made up almost entirely of $(H^1H^1H^1)^+$. Then a mixture of the two samples was admitted at such a pressure that $(H^1H^1H^1)^+$ and $(H^1H^2)^+$ were present in equal numbers and again the shape of the peak was examined. If the two ions are of different weight one would expect the peak in the latter case to appear somewhat broadened. No such broadening could be detected. It is estimated from these measurements that the packing fraction of H^2 lies between 4×10^{-3} and 12×10^{-3},

$$\text{or } 2.024 > W > 2.008$$

where W represents the atomic weight.*

It may be well to point out that all the results of this paper simply point to the conclusion that a small number of normal molecules of mass 3 does exist in hydrogen. Conrad[9] has found evidence for the formation of neutral H_3 molecules in a discharge tube but it has been shown[10] on quantum mechanical grounds that three normal H atoms cannot combine to form a stable triatomic molecule. Hence if a neutral triatomic molecule exists it must be in an excited state in which case the ion would be formed as the result of a secondary or tertiary process and would not be a linear function of the pressure. Even without considering the weight of the spectroscopic evidence[11] the most reasonable assumption which accounts for the results of this experiment is the presence of H^1H^2 molecules in hydrogen.

Results for Helium

Helium purified in a misch-metal arc was admitted to the tube and a search made for isotopes 3 and 5. None were found. It is believed that a concentration of one in 50,000 or less could have been detected. This negative result agrees with that of Kallmann and Lasareff[8] but is a test not as discriminating as theirs.

The author is indebted to Professor Urey and Drs. Brickwedde and Murphy for the samples of hydrogen which made this experiment possible. He is also grateful to the members of the Palmer Laboratory staff for their kindly interest and timely suggestions.

* The author has just recently learned through private communication that Dr. K. T. Bainbridge has succeeded in measuring the atomic weight of H^2 with his mass spectrograph. His result is $W = 2.0135$.

[9] R. Conrad, Zeits. f. Physik **75**, 504 (1932).

[10] W. Heitler and F. London, Zeits. f. Physik **44**, 455 (1927).

[11] The existence of the $H_\alpha{}^2$ line has been confirmed by Shenstone and Turner using a grating in this laboratory and some of the "sample III" hydrogen.

Elastic and Inelastic Electron Scattering in Hydrogen

By A. L. HUGHES* *and* J. H. MCMILLEN
Washington University, St. Louis

(Received May 16, 1932)

A study of both the *energy distribution and the angular distribution of electrons scattered by hydrogen molecules* was made. *Elastic scattering* was investigated in the range between 40° and 165°, for collision energies between 35 and 200 volts. The number of electrons scattered at any angle was smaller the greater the energy of impact. In all the curves a minimum was found between 90° and 110°, while for the 35 and 50 volt collisions the curves rose to a definite maximum at 155° to 160°. *Inelastic scattering* was found to occur with energy losses ranging from a minimum of about 12.0 volts, through a sharply defined *most probable* loss at 12.6 volts, to a complete loss of all the energy. It was found that as the energy of collision was increased the probability of the larger energy losses increased with respect to the probability of the smaller energy losses. Angular distribution curves were steeper, the smaller the amount of energy lost at collision for a given collision energy. Also, for any particular energy loss, the steepness increased with the speed of the colliding electrons. The electrons which are torn off from atoms in ionizing collisions are termed *ejected electrons* and are presumably those which appear in our experiment with but little or no energy. The angular distribution of the ejected electrons, having various amounts of energy (1 to 8 volts), were studied for different collision energies (35 to 340 volts). In every case there was a noteworthy absence of considerable scattering at small angles, so characteristic of the scattering of faster electrons. In many cases, pronounced peaks were found in the angular distribution curves at large angles (90° to 160°). These peaks change in position and size in a regular manner with the energy of collision and the energy of the ejected electron.

THIS paper is an account of a continuation of our experimental investigations on electron scattering in gases. The results of our investigations on argon have already been published in this journal.[1] As argon is monatomic, it seemed desirable to select a diatomic gas for the second investigation; hence hydrogen was chosen for study. The method of experimentation with hydrogen was identical with that used in the work on argon.

Scattering of electrons by hydrogen has been studied by Harnwell,[2] McMillen,[3] Arnot,[4] Bullard and Massey,[5] and by Ramsauer and Kollath.[6] Harnwell and McMillen found that the number of electrons scattered elastically fell off steeply as the angle was increased (angle range 7° to 60°; energy range 40 to 180 volts). Arnot, working with electrons between 30 and

* This work was made possible by assistance to the senior author from a grant made by the Rockefeller Foundation to Washington University for research in science.

[1] A. L. Hughes and J. H. McMillen, Phys. Rev. **39,** 585 (1932). It is convenient to mention here that several paragraphs were accidentally omitted from the end of the paper referred to. The omission was rectified in a letter to the Editor of the Physical Review **40,** 469 (1932). An error in Eq. (2) of the same paper was also corrected.

[2] G. P. Harnwell, Phys. Rev. **33,** 559 (1929); **34,** 661 (1929); **35,** 285 (1930).

[3] J. H. McMillen, Phys. Rev. **36,** 1034 (1930).

[4] F. L. Arnot, Proc. Roy. Soc. **133,** 615 (1931).

[5] E. C. Bullard and H. S. W. Massey, Proc. Roy. Soc. **133,** 637 (1931).

[6] C. Ramsauer and R. Kollath, Ann. d. Physik **12,** 529 (1932).

800 volts energy, confirmed these results, but found that, at about 90°, the curve representing the number of elastically scattered electrons began to increase. (His angular range extended to 120°.) Similar results were obtained by Bullard and Massey who studied the elastic scattering of electrons having energies between 4 and 30 volts. Ramsauer and Kollath also found, for very slow electron (1.5 to 11.0 volts), a definite but rather slow increase in the elastic scattering as the angle was increased above 90° to as far as 165°. So far as we know there have been no investigations of the distributions of electrons scattered inelastically by hydrogen.

Experimental Method

The same apparatus was used in the study of electron scattering in hydrogen as in the previous work on argon.[7] The reader is therefore referred to the previous article for a description of the apparatus and method of experimentation. Hydrogen was prepared by the action of pure dilute sulphuric acid on pure zinc. The gas was dried by passing through a tube immersed in liquid air. It was then stored in a reservoir over mercury and admitted to the apparatus through a capillary tube and a liquid air trap to freeze out the mercury. A suitable constant pressure could be maintained in the apparatus by regulating the driving pressure of the hydrogen on the high pressure side of the capillary tube, the gas being pumped out of the apparatus by a condensation pump. This constant flow method provides us with pure hydrogen in the scattering experiments. In experiments on scattering in hydrogen it is particularly important to avoid impurities, as the scattering by a hydrogen molecule is small compared with the scattering by heavier molecules.

Our results on hydrogen are expressed in the same units as those used in the previous article; that is, we measure the number of electrons diverted from the main beam at an angle θ and then multiply by $\sin\theta$ to compensate for the change with angle of the effective scattering path length. The final values are therefore proportional to the scattering coefficient as defined on page 589 of our previous paper.[8]

Elastic Scattering

In view of the recent comprehensive work of Bullard and Massey and of Arnot on elastic scattering of electrons in hydrogen, we did not make an extensive study of the elastic scattering. It seemed worth while, however, to attempt to extend the studies to large angles, inasmuch as the investigators referred to did not go beyond about 120°. Our results are shown in Fig. 1. It will be seen that we confirm the discovery of Bullard and Massey and of Arnot that there is a minimum at about 90°. In addition, we find a definite maximum at large angles which is contrary to the expectations of Arnot who thought that "the curves probably continue to rise steadily as the angle is increased to 180°."

We have indicated the positions of Arnot's 29 volt curve and Bullard

[7] A. L. Hughes and J. H. McMillen, Phys. Rev. **39**, 585 (1932).
[8] A. L. Hughes and J. H. McMillen, Phys. Rev. **39**, 585 (1932).

and Massey's 30 volts curve in the neighborhood of our 35 volt curve. It is evident that although all three curves agree in giving a minimum at about 90°, no two curves are in satisfactory accord. We can only say, therefore, that there is qualitative agreement insofar as the curves rise and fall together.

Inelastic Scattering

The distributions of energy lost by 50 volt primary electrons, when scattered at various angles, are shown in Fig. 2. The various curves have been arbitrarily adjusted to the same height for a loss of energy amounting to 20 volts. It will be seen that the "most probable" energy loss is approximately 12.6 volts. It is evident that the larger energy losses are relatively more

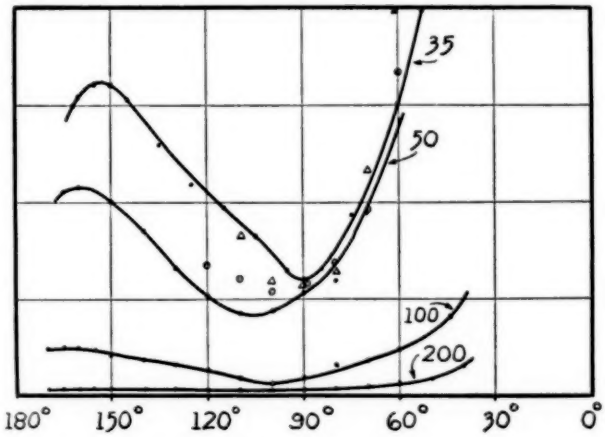

Fig. 1. Elastic scattering of electrons by hydrogen molecules. The numbers attached to the curves represent the electron energies in volts. Circles with dots, Arnot's curve for 29 volt electrons. Triangles, Bullard and Massey's curve for 30 volt electrons.

probable for the larger scattering angles. A small subsidiary maximum appears at about 16 volts in the 30° and 45° curves. In Fig. 3 we have the energy distributions of electrons, all scattered at 10°, for various primary energies between 35 and 200 volts. (The curve for 412 volt electrons, scattered at 5°, is included). As before, the curves are arbitrarily fitted together at the 20 volt loss. It is seen that, for a given scattering angle, here 10°, the 12.6 volt loss becomes relatively *more* probable the *lower* the primary electron energy. It will be noticed that there is a distinct slowing up in the rate of the fall of the curve between 15 and 20 volts, and in one curve (the 412 volt, 5° curve) there is a small subsidiary maximum at 17.5 volts.

Various energy losses, *viz.*, 12.6, 16.4, and 25.0 volts, were selected for a study of their angular distributions. The 12.6 volt loss, the most probable loss, is an excitation loss. The 16.4 loss is just above the ionization potential which is 15.9 volts. (The 15.9 volt loss was not studied, as the resolution of the apparatus was imperfect and consequently the electrons selected at a 15.9 volt setting would really include a number of electrons causing excitation). Finally the 25 volt loss was chosen because it was well above the minimum ionization loss and, at the same time, the yield was not too small to

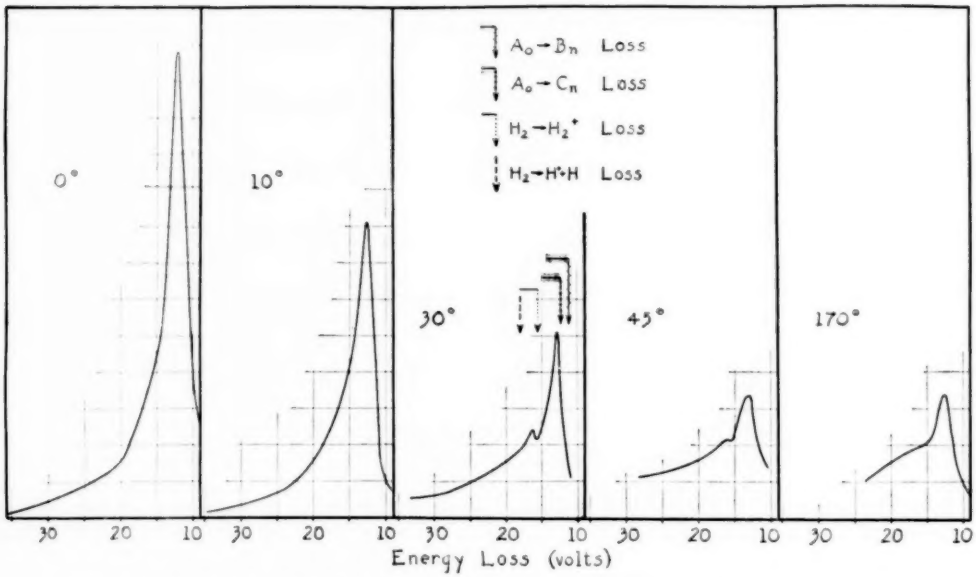

Fig. 2. Energy distribution of electrons scattered inelastically by hydrogen molecules, showing variation with *angle*. Primary energy in all cases = 50 volts. All curves are adjusted arbitrarily to the same height at the 20 volt loss. On the 30° curve, the positions of the theoretically possible types of effects are indicated by suitable vertical lines. The associated horizontal lines indicate the possible vibrational levels.

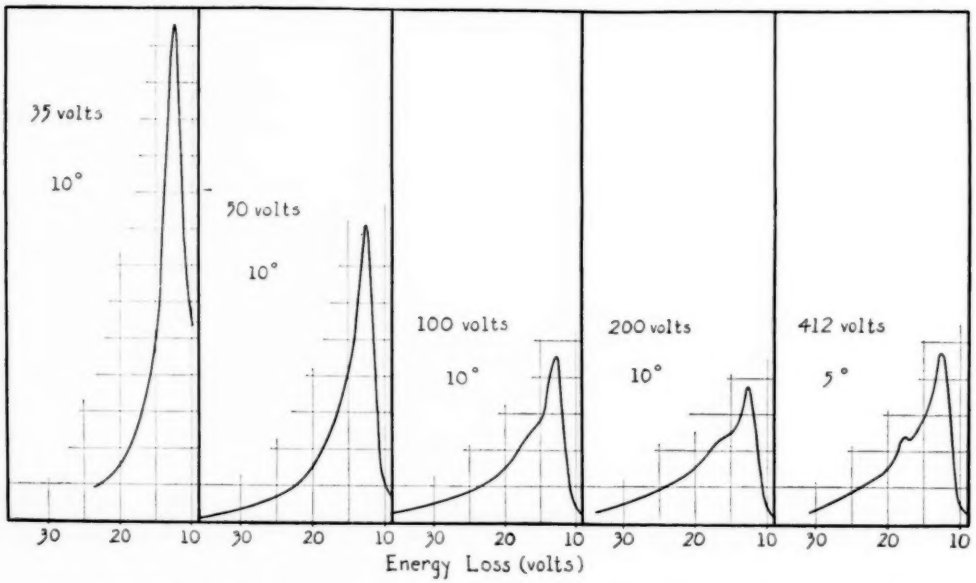

Fig. 3. Energy distributions of electrons scattered inelastically by hydrogen molecules, showing variation with *primary energy*. All curves are adjusted arbitrarily to the same height at the 20 volt loss.

make measurements unreliable. According to Fig. 4, the less the energy loss, the steeper the angular distribution curve, a result already implied in the previous paragraph. The curves are all adjusted to the same value at 10°.

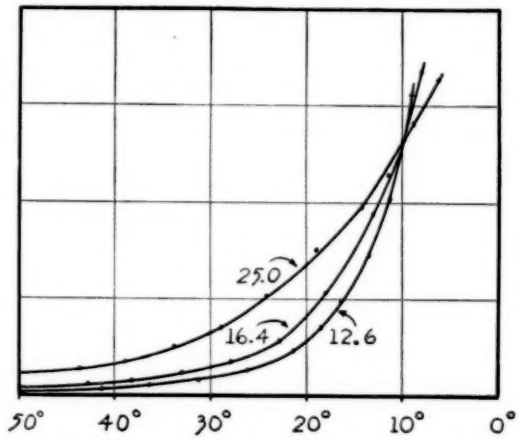

Fig. 4. Angular distribution of electrons scattered inelastically. Primary electron energy in all cases = 50 volts. Energy losses in volts are indicated on the curves. All curves are adjusted arbitrarily to the same height at 10°.

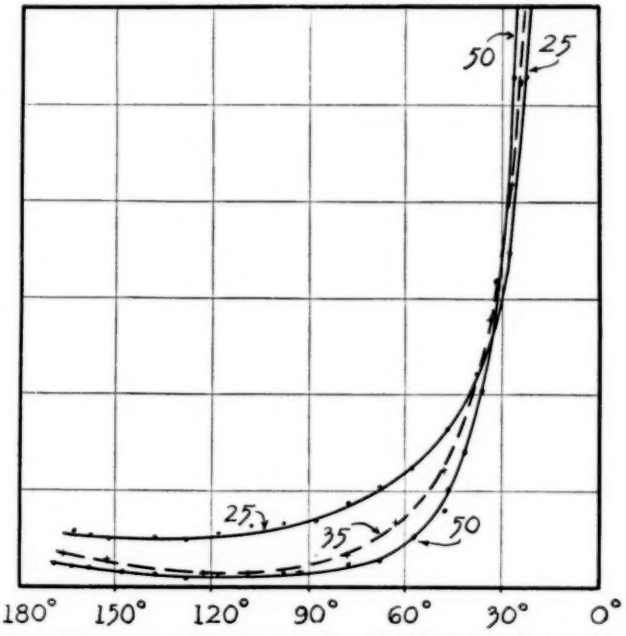

Fig. 5. Angular distribution of electrons scattered with 12.6 volts loss of energy, showing dependence on primary electron energy. The primary energy, expressed in volts, is indicated for each curve.

Finally, in Fig. 5, we show how the angular distribution curves for a constant loss, *viz.*, 12.6 volt (corresponding to excitation of the molecule), depend on the energy of the electrons before collision. Up to about 90° or so, the curves are steeper the higher the primary energy. (Too much reliance

should not be placed on the exact shape of the curves above 90°, as the readings were extremely small).

Ejected Electrons

Following the terminology used in our previous paper on argon,[9] we call those electrons which appear with very low energies the ejected electrons. When ionization takes place, an electron is detached from the atom. Consequently, after an ionizing collision we have two electrons; the faster we assume to be the colliding electron and the slower the ejected electron. (We

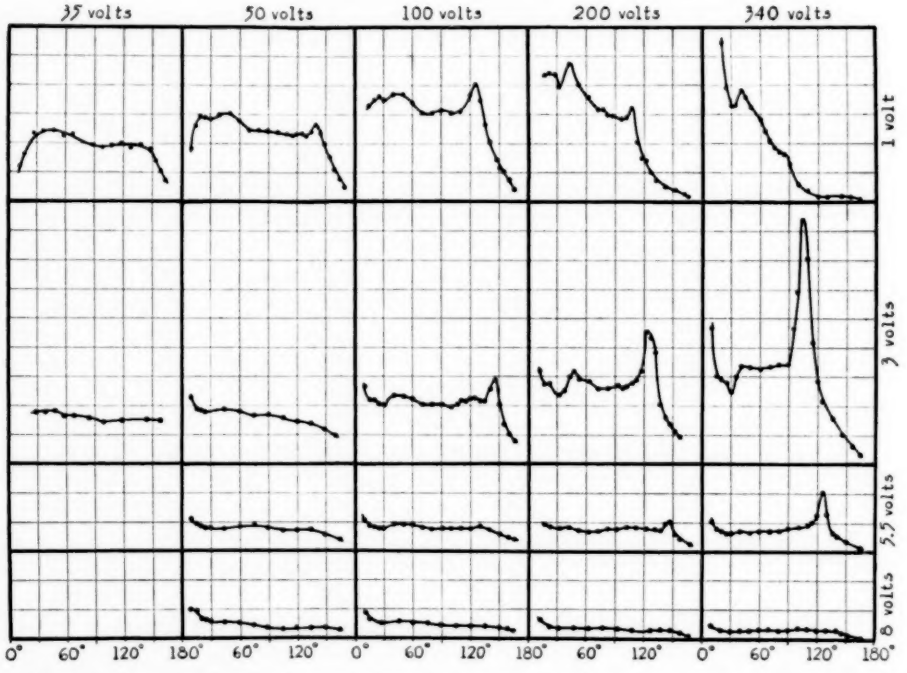

Fig. 6. Angular distributions of ejected electrons. Each row represents a single ejected electron energy which takes the values 1, 3, 5.5, and 8 volts as we go from one row to the next. Each column corresponds to a different energy of collision. All curves are to the *same* scale.

cannot, of course, demonstrate that the slower of the two electrons is the one ejected from the atom; it seems probable, however, and suggests a convenient terminology.) For a study of the angular distribution of ejected electrons we selected the 1, 3, 5.5, and 8 volt ejected electrons resulting from a collision between hydrogen molecules and primary electrons having energies ranging from 35 to 340 volts. The results are shown in Fig. 6. Perhaps the most striking feature common to all the curves is the absence of the excessively large scattering at small angles which is so characteristic of electrons scattered inelastically with considerable amounts of energy (see Fig. 4). The next feature worthy of mention is the presence, in many cases, of certain highly preferred directions for the emission of ejected electrons. A certain regularity in the way in which these peaks shift can be traced. Thus for any

[9] A. L. Hughes and J. H. McMillen, Phys. Rev. **39**, 585 (1932).

one ejected electron energy, an *increase* in the energy of the primary electrons is accompanied by a progressive shift of the position of the peak to *smaller* angles. On the other hand, if we keep the primary energy fixed and study the various groups of ejected electrons, we find that the *higher* the energy of the ejected electrons the *larger* the angle at which the peak is to be found. For 8 volt ejected electrons the angular distribution curves have become practically featureless. It is to be noticed, too, that the peaks in the angular distribution curves for all the ejected electrons practically disappear when the primary energy is low (35 and 50 volts).

Discussion

Several theoretical physicists have amplified and extended the pioneer work of Born, who first applied wave mechanics to the scattering of electrons by atoms, with the result that certain features of electron scattering are now accounted for theoretically. Very little, however, has been done on the corresponding problem for molecules. Massey and Mohr[10] have carried through a calculation of the scattering of electrons by diatomic molecules. It is assumed that the diatomic molecule may be represented by a suitable potential field having axial symmetry over which passes an electron wave. The amplitude of the wave scattered in any desired direction is calculated. An average is then taken for all possible orientations of the molecular axis with respect to the direction of the primary wave. Finally it is found that the scattering by such a diatomic molecule is obtained on multiplying what is substantially the scattering by the constituent atoms by a diffraction factor $(1+\sin x/x)$, in which $x=4\pi d \sin(\theta/2)/\lambda$, where d is the distance between the atoms in the molecule and λ the electron wave-length. This diffraction factor changes sign periodically as x increases. Because of the extremely steep character of the curve for atomic scattering, however, it is not easy to detect the superposed oscillation of the diffraction pattern.[11] Thus Massey and Mohr were unable to show that Arnot's experimental curves (for 80 volts and above) for the scattering of electrons by hydrogen were in better accord with their formula for scattering by molecular hydrogen than with the formula for scattering by atomic hydrogen. Massey and Mohr state that the scattering curves obtained by McMillen[12] fall off much more rapidly with angle of scattering than the calculated values. This is incorrect, for McMillen's curves are almost superposable on Arnot's, and if anything, they are less steep. Arnot's range of angles was nearly twice as great as McMillen's, and so afforded a better test.

The rise in the scattering curves with angle as the angle is increased beyond about 90° as shown in Fig. 1 of this paper and in Arnot's work is not accounted for by Massey and Mohr's theory. This theory is probably too simplified to do more than describe the part of the scattering curves below

[10] H. S. W. Massey, Proc. Roy. Soc. **129,** 616 (1930); H. S. W. Massey and C. B. O. Mohr, Proc. Roy. Soc. **135,** 258 (1932).

[11] The oscillations of intensity superposed on rapidly falling atomic scattering curves, contributed by factors such as $(1+\sin x/x)$ have been considered by Mark and Wierl, principally in connection with their scattering of high speed electrons. (See H. Mark and R. Wierl "Die experimentellen und theoretischen Grundlagen der Electronenbeugung." Gebrüder Borntraeger, Berlin, 1931).

[12] J. H. McMillen, Phys. Rev. **36,** 1034 (1931).

about 90°. To account qualitatively for the presence of maxima and minima in the scattering of electrons by *atoms*, Bullard and Massey[13] called attention to the necessity for considering the effects of distortion of the electron wave, electron exchange and polarization. These effects have not been included in Massey and Mohr's treatment of scattering by diatomic molecules.

We are accustomed to think of a clear cut distinction between an inelastic and an elastic collision between electrons and atoms. In the latter case the electron leaves the atom unchanged, while in the former case the electron excites (or ionizes) the atom. The lowest excitation state differs in energy so

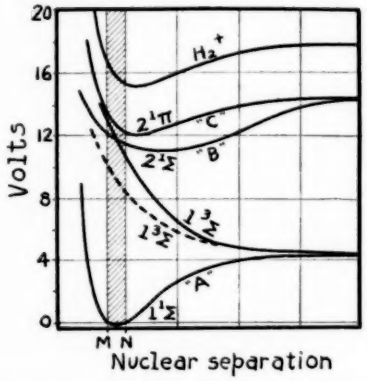

Fig. 7. Franck-Condon diagram for hydrogen. The diagrams given by Smyth and by Mulliken agree substantially in the case of all curves except the $1^3\Sigma$ curve. Mulliken's $1^3\Sigma$ curve is shown as a broken line.

considerably from the normal state that there is no difficulty in identifying an electron which has rebounded from an atom elastically and one which has rebounded inelastically. With a molecule, however, it is conceivable that an electron collision could change the vibrational state and the rotational state of the molecule without altering the electronic state. We should therefore be prepared, perhaps, to find evidences of energy losses of the order of a few tenths of a volt. Experimentally these might cause an asymmetrical broadening of the so-called elastic peak towards the low energy side. We have made as yet no special search for such an effect but it seems possible that it could be detected in view of the experimental work of Ramien[14] on the average energy lost by electrons in colliding with many hydrogen molecules, and in view of the theoretical discussion of Massey.[15]

The possible molecular changes resulting in a hydrogen molecule after an electron impact may be visualized by the help of a Franck-Condon diagram. In Fig. 7 we reproduce a somewhat simplified copy of the diagram given by Smyth[16] for hydrogen. The curve A is for the normal molecule, and the stable state is the one corresponding to a nuclear separation varying between M and N (corresponding to the lowest vibrational state). Excitation by electron impact is represented by a vertical displacement, within the shaded area, to one of the other curves. If the transition is to the $1^3\Sigma$ state, the atoms immediately fly apart and dissociation results. If, on the other hand, the transi-

[13] E. C. Bullard and H. S. W. Massey, Proc. Roy. Soc. **A133,** 637 (1931).
[14] H. Ramien, Zeits. f. Physik **70,** 353 (1931).
[15] H. S. W. Massey, Proc. Camb. Phil. Soc. **28,** 99 (1932).
[16] H. D. Smyth, Rev. Mod. Physics **3,** 347 (1931).

tion is to the B state (i.e., to the $2^1\Sigma$ state), we get an excited molecule which does not dissociate so long as it is in this state. The energy necessary to produce either state is not sharply defined because of the finite width of the shaded column resting on MN. Thus a transition to the $1^3\Sigma$ state requires anything between 10.9 volts and 13.5 volts, while a transition to the $2^1\Sigma$ state requires an amount of energy between 11.5 and 12.2 volts and a transition to the $2^1\Pi$ state calls for something between 12.2 and 13.6 volts. This overlapping makes it impossible to decide whether our most probable loss—12.6 volts—corresponds to a transition to the $1^3\Sigma$ state to the $2^1\Sigma$ state or to the $2^1\Pi$ state. It should be stated that the Franck-Condon diagram as given by Mulliken[17] differs slightly in scale from that given by Smyth, yet in such a way as to suggest a different interpretation of our results. Briefly, if the Mulliken representation is correct, we cannot regard our 12.6 volt loss as indicating transitions to either the $1^3\Sigma$ or to a low vibrational level of the $2^1\Sigma$ state; we must attribute it to a transition to the $2^1\Pi$ state, or to a high vibrational level of the $2^1\Sigma$ state. Dr. E. U. Condon informs us that, on account of the approximations made in the calculations of these curves, they cannot be regarded as exact; thus the position of the $1^3\Sigma$ curve above MN (Fig. 7) may be in error by 2 or 3 volts. In view of this situation it is evident that the identification of our 12.6 volt loss with one or other of the possible transitions cannot be made with certainty. It may well be that our 12.6 volt peak corresponds to transitions to all three states. In view of the sharpness of the energy losses to be expected in *atomic* excitation, as compared with this blurring to be expected in molecular excitation, it is possible that the most probable energy loss peak in hydrogen (e.g., in Fig. 2) would be appreciably less steep (especially on the side next to the elastic peak) than the 11.6 volt peak in argon taken under identical conditions. Tests, however, showed that if the 12.6 volt peak in hydrogen and the 11.6 volt peak in argon were superposed, then the hydrogen peak had no more *spread* toward the elastic peak than had the argon peak.

Massey and Mohr have calculated by wave mechanics the angular distribution of electrons which collide inelastically with hydrogen molecules to give rise to the $2^1\Sigma$ state. These distributions have maxima at certain angles, e.g., for 30 volt primary electrons, a sharp maximum appears at 20°. Fig. 5 gives the angular distribution of the 12.6 volt loss. Unfortunately the curve does not extend to angles smaller than 25°, but it seems obvious that there is no indication of the presence of a maximum at 20°. Massey and Mohr do not give curves for the angular distributions of electrons which have collided with hydrogen molecules and put them into the $1^3\Sigma$ state. Consequently we are unable to compare them with our experimental curves. There is one point of significance, however. By integrating over all angles Massey and Mohr have determined the probabilities of excitation of the $2^1\Sigma$ state and the $1^3\Sigma$ state of hydrogen as a function of the energy of electron impact. The scale of their curves implies that the probability of excitation of the $1^3\Sigma$ transition is about 10,000 times as large as the probability of excitation of the $2^1\Sigma$ state. If this is correct, then our 12.6 volt peak must be associated almost wholly with transitions to the $1^3\Sigma$ state and the number of transitions to the

[17] R. S. Mulliken, Rev. Mod. Phys. **4,** 1 (1932); see also W. Finkenburg and W. Weizel, Zeits. f. Physik **68,** 577 (1931).

$2^1\Sigma$ state must be neglibly small. (This ignores, of course, the possibility of transitions to the $2^1\Pi$ state).

In the case of scattering of electrons by argon atoms, we saw that if the energy lost by the colliding electrons was greater than 15.7 volts, the ionization potential of argon, then a second electron, which we termed the ejected electron, must appear. Thus the parts of the scattering curves corresponding to energy losses greater than 15.7 volts give us the distributions of the colliding, or ionizing electrons, as well as the distributions of the ejected electron. It was found convenient to assume that the ejected electron was always the slower of the two electrons moving away from the atom after an ionizing collision. The situation is still more complicated when we have a diatomic molecule like hydrogen instead of an atom like argon. When the collision is of the type which leaves a molecular ion, we have, as in the case of argon, two electrons and an ion. However, as Tate, Lozier and Bleakney[18] have shown, some types of ionization result in dissociation with the constituent atoms flying apart with considerable kinetic energy. We thus have, after collision, two ions and two electrons moving away from the place of collision, each with its own angle distribution characteristic of its energy and of the energy of the electron before collision. So far as we know, no attempt has been made to solve such a problem theoretically. It is evident, from the curves shown in Fig. 6, that, when electrons are ejected with a certain amount of energy from certain types of collisions, strongly preferred directions of emission are to be found. We can also trace certain regularities in the displacement of the peak representing preferred directions of emission as the energies concerned are progressively altered. Thus, for a given energy of ejection, the position of the peak moves to smaller angles as the energy of collision is increased. On the other hand, if we keep the energy of collision unchanged we find that, as the energy of the ejected electron is increased the position of the peak moves to larger angles. There is also a regularity in the way in which the magnitude of the peak changes as we go across a row or down a column. A comparison of these results with those for argon will show that the distributions of electrons ejected from argon atoms are similar in certain respects to the distributions for hydrogen.

The results for hydrogen were obtained with the same apparatus as that in our work on argon. In our previous article we pointed out that the method of speeding up the electrons by a suitable field into the analyser introduced a certain amount of *blurring* of the sharpness of the curves which was greater the smaller the energy of the electrons under investigation. The same comment applies also to this work on hydrogen.

So far as we know, there has been no attempt to work out a theory for the distributions of the ejected electrons. The frequent occurrence of strongly preferred directions of emission for these electrons suggests a peculiar kind of diffraction effect.

We take great pleasure in thanking Dr. E. U. Condon for his helpful comments on the situation relating to the interpretation of the Franck-Condon diagram for hydrogen.

[18] W. Bleakney, Phys. Rev. **35**, 1180 (1930); W. W. Lozier, Phys. Rev. **36**, 1285 (1930); J. T. Tate and W. W. Lozier, Phys. Rev. **39**, 254 (1932).

Electronic Structures of Polyatomic Molecules and Valence. II. General Considerations

By ROBERT S. MULLIKEN
Ryerson Physical Laboratory, University of Chicago

(Received April 23, 1932)

A description of electronic structures of polyatomic molecules in terms of atomic and molecular orbitals (one-electron orbital wave functions) is suggested, as follows. Unshared electrons are considered as occupying atomic orbitals associated with the various nuclei to which the electrons belong. Shared electrons are described from several points of view, whose simultaneous consideration should give a better understanding of their condition and functions in the molecule. In the first place, they are described, in terms of atomic orbitals, from the point of view of *each nucleus* in the molecule. In the second place, they are described from a unitary molecular point of view, in terms of molecular orbitals. It is pointed out that the chemical evidence which led Lewis to his concept of the electron-pair bond can now all be explained by the quantum theory without the necessity of using such a concept. It is noted that the Heitler-London, Pauling-Slater quantum-mechanical electron-pair bond concept differs markedly from Lewis's, also that it is of more restricted application to chemical data. It is pointed out, in agreement with Hund, that properties of the H. and L., S. and P. electron-pair bond which make it useful in dealing with chemical combination are also possessed by the concept of molecular orbitals. For example, the Pauling-Slater criterion that bonds which correspond to a maximum overlapping of atomic orbitals are the strongest is just as characteristic of bonding molecular orbitals as of electron-pair bonds. An electron-pair bond is here interpreted as being little other than two electrons occupying a bonding molecular orbital. Or in general one must say that a set of n electron-pair bonds is interpreted as a set of $2n$ electrons occupying n molecular orbitals, because molecular orbitals are not necessarily localized *between two nuclei* like electron-pair bonds. It is concluded that the essential facts of molecular electronic structure can be qualitatively understood in terms of the mode of description stated in the first paragraph above and of a semi-empirical valence rule (not essentially new) which summarizes, in the light of quantum theory, the most important regularities in regard to the types of chemical compounds which are stable.

ATOMIC AND MOLECULAR ORBITALS FOR DESCRIBING MOLECULAR ELECTRONIC STRUCTURES

1. Introduction

THE series of which the present paper is the second[1] is a development of a program initiated by Hund[2] and the writer[3] and carried forward principally by us, by Herzberg, and by Dunkel.[4,5,6,7,8,9,10,11,21] The object

[1] A brief preliminary paper outlining some of the principles used and some of the main results has already been published: R. S. Mulliken, Phys. Rev. **40,** 55 (1932). In this paper, the following changes should be made in the last paragraph on p. 58. Line 8, delete "tetrahedralized but uncombined"; lines 9 and 12, delete brackets around σ, and delete "uncombined" before $2p\sigma$; lines 10–11, delete "tetrahedralized uncombined." Also on p. 59, line 10, delete brackets. Elsewhere, do not delete brackets or "tetrahedralized."

[2] F. Hund, Zeits. f. Physik **51,** 759 (1928).

[3] R. S. Mulliken, Phys. Rev. **32,** 186, 761 (1928); **33,** 730 (1929).

of this program has been to describe and understand molecules in terms of one-electron orbital wave functions of distinctly molecular character. The first paper of the present series[1] will be referred to hereafter as I. The present paper is to a large extent a preliminary critical review rather than a presentation of new results.

2. Atomic and molecular orbitals

From here on, one-electron orbital wave functions will be referred to for brevity as *orbitals*. The method followed here will be to describe unshared electrons always in terms of *atomic orbitals* but to use *molecular orbitals* for shared electrons. This procedure was first used by Lennard-Jones[12] for diatomic molecules, and is a partial departure from the original formulation of the above program. Among the main objects of the present series are the determination of the forms of molecular orbitals and the application of a knowledge of these forms to an explanation of valence and related chemical phenomena.

Shared electrons include two types, namely *bonding* and *anti-bonding* electrons, at least in diatomic molecules; unshared electrons belong to the class of *non-bonding* electrons.[8,9,10,11] Anti-bonding electrons occur in diatomic molecules only when accompanied by a larger number of bonding electrons. Often it is hard to draw the line between shared and unshared electrons; in such cases, the electrons in question consist of bonding and anti-bonding electrons in equal numbers, at least in diatomic molecules (cf. third following paragraph and references 8, 9).

By an atomic orbital is meant an orbital corresponding to the motion of an electron in the field of a single nucleus plus other electrons, while a molecular orbital corresponds to the motion of an electron in the field of two or more nuclei plus other electrons. Both atomic and molecular orbitals may be thought of as defined in accordance with the Hartree method of the self-consistent field, in order to allow so far as possible for the effects of other electrons than the one whose orbital is under consideration.

Every non-degenerate orbital can be occupied by at most two electrons,

---

[4] F. Hund, Zeits. f. Physik **63**, 719 (1930).

[5] F. Hund, Zeits. f. Physik **73**, 1 (1931); **74**, 429 (1932).

[6] F. Hund, Zeits. f. Physik **73**, 565 (1932).

[7] F. Hund, Zeits. f. Physik **74**, 1 (1932).

[8] R. S. Mulliken, Chem. Rev. **9**, 347 (1931). On p. 351 (foot), p. 353 (middle) and p. 355 (middle), the writer gives the incorrect impression that, omitting the energy of repulsion of the nuclei, the promoted wave-function $2p\sigma$ has a *higher* energy than the original $1s$ atomic wave-function. Actually the energy is a little lower (for $R>0$) than that of $1s$ in the H atom. This, however, does not in any essential way affect the validity of the arguments given. The fact that the energy of $2p\sigma$ goes up sufficiently rapidly *with reference* to that of $1s\sigma$, in particular the fact that the energy of $2p\sigma$ plus the nuclear repulsion energy causes a net repulsion of H and H^+, is sufficient.

[9] R. S. Mulliken, Rev. Mod. Phys. **4**, 1 (1932).

[10] G. Herzberg, Zeits. f. Physik **57**, 601 (1929).

[11] G. Herzberg, "Molekülstruktur," Leipziger Vorträge, 1931, p. 167. S. Hirzel, Leipzig.

[12] J. E. Lennard-Jones, Trans. Faraday Soc. **25**, 668 (1929).

corresponding to the two possible orientations of the electron spin. Orbitals having n-fold degeneracy can be occupied by at most $2n$ electrons. Examples of non-degenerate orbitals are s orbitals in atoms, σ orbitals in diatomic or linear molecules[9,11] and most orbitals in non-linear polyatomic molecules. Examples of degenerate orbitals are p, d, . . . orbitals in atoms ($2l+1$-fold degeneracy), π, δ, . . . orbitals in diatomic or linear molecules (2-fold degeneracy), $[\pi]$ orbitals (2-fold degeneracy), in molecules like NH_3, NO_3^-, etc. (cf. I), d_γ (2-fold degeneracy) d_ϵ and p (3-fold degeneracy) in molecules having tetrahedral or octahedral symmetry (cf. I).

It should be noted that the use of atomic orbitals for describing the condition of unshared electrons in molecules often gives a false impression of the amount of degeneracy. For example if one writes $1s^2 1s^2\sigma 2s^2$ for the electron configuration[9] of Li_2, the description suggests only one ionization energy for the $1s$ electrons, since $1s$ orbitals on two Li atoms are identical in energy. If one writes $\sigma 1s^2\sigma^* 1s^2\sigma 2s^2$, however, using molecular orbitals exclusively,[9] it is made explicit that there are two distinct energies corresponding to $\sigma 1s$ and $\sigma^* 1s$. In this particular example the energy difference is negligible, but in many cases, especially where there are electrons which are near the border line between unshared and shared, considerable energy-splittings may exist.

For example in O_2 one may perhaps best write $1s^2 1s^2 2s^2 2s^2$ . . . but must then grant that the $2s$ type, although essentially unshared, has a markedly double ionization energy.[9] In N_2, the sharing of the $2s$ electrons is so strong that it is best to write $1s^2 1s^2\sigma 2s^2\sigma^* 2s^2$, the types $\sigma 2s$ and $\sigma^* 2s$ being very different in energy.[9] In CF_4, one might write $\{1s^2 2s^2 2p[\pi]^4\}_F{}^4\{1s^2\}_C a^2 b^6$, where a and b represent the shared electrons. In so doing, it should be recognized that four distinct energy values are to be expected for the unshared $1s$ fluorine electrons, likewise for the $2s$ and for the $2p[\pi]$,—or possibly eight values for $2p[\pi]$ because of a possible splitting up of the degeneracy implied by the symbol $[\pi]$. The four-fold splitting might be appreciable for the $2p[\pi]$ and possibly also for the $2s$, although of course negligible for the $1s$. The carbon $1s$ is of course very different in energy from the fluorine $1s$. These examples should be sufficient to show how to guard against a possible misinterpretation of the use of atomic orbitals in describing molecules.

In the present method, molecular orbitals are conceived of as entities quite independent of atomic orbitals. Nevertheless in practise molecular orbitals can usually be conveniently approximated by building up linear combinations of orbitals of the atomic type. The present method of thinking in terms of the finished molecule, used already by Lewis in his valence theory, avoids the disputes and ambiguities, or the necessity of using complicated linear combinations, which arise if one thinks of molecules as composed of definite atoms or ions.

3. One-nucleus and other viewpoints

Understanding of the electronic structure of molecules is greatly aided by introducing a set of partially overlapping descriptions whereby the electrons immediately surrounding each nucleus in a molecule are described in

terms of a set of atomic orbitals corresponding to the viewpoint of that nucleus. (Sidgwick,—reference 13, p. 98,—has used essentially this viewpoint.) Such *one-nucleus viewpoint* atomic orbitals must be taken as considerably deformed compared with ordinary atomic orbitals, the deformation being thought of as caused by the strong fields of the other nuclei. Unshared electrons, of course, are here assigned exclusively each to a particular nucleus, so that the one-nucleus viewpoints for them give the same description as that introduced in the first paragraph of section 2. Shared electrons, however, being considered as belonging to two or more nuclei, receive non-mutually-exclusive descriptions in terms of deformed atomic orbitals associated with each, or at least with several, of these nuclei (cf. descriptions of N_2^+, HCl, and $PtCl_6^=$ in section 8 below and of CF_4 and other molecules in I as examples).

An understanding of molecular electronic structures can not infrequently be advanced by using still other auxiliary viewpoints. The well-known *united-atom* viewpoint is sometimes useful (cf. reference 9, p. 19). In hydrides especially, this or, rather, a viewpoint in which each hydrogen nucleus is at first regarded as united with a neighboring nucleus of larger charge, is valuable. This amounts to saying that in hydrogen compounds, the hydrogen one-nucleus viewpoints can often appropriately be treated as unimportant compared with the viewpoints of other nuclei in the molecule or that the hydrogen nuclei can be regarded merely as perturbing force-centers. The validity of this point of view is indicated by a number of facts concerning diatomic hydride band spectra,[9] also by chemical data, notably the behavior of the boron hydrides.[8]

The familiar chemical method of regarding many molecules as built up of "radicals" suggests the usefulness of a point of view making use of shared orbitals belonging to *radicals* instead of, or in addition to, the viewpoint of molecular shared orbitals.

Quantum Theory of Valence

4. Review of theories of valence and molecular structure

A semi-historical survey of some of the theories dealing with valence and molecular structure will put the whole problem in better perspective. Incidentally, the writer hopes to show that there are no compelling reasons, either empirical or theoretical, for placing primary emphasis on electron pairs in constructing theories of valence. He hopes thereby to remove possible objections to the present method based on its lack of such emphasis.

The best chemical theory of valence covering all types of compounds is generally agreed to be that developed principally by G. N. Lewis.[13] To a rather large extent, the essential features of this theory still stand, although their meaning has been made clearer and more specific by interpreting them in the light of the quantum theory. The most important features of Lewis's theory are perhaps the following:

[13] G. N. Lewis, Valence and The Structure of Atoms and Molecules. The Chemical Catalog Co., New York, 1923. Sidgwick's excellent book should also be consulted: N. V. Sidgwick, The Electronic Theory of Valency. The Clarendon Press, Oxford, 1927.

(1)(*a*) Every atom tends so to give, take, or share electrons as to be surrounded by an outer "group of eight" electrons (reference 13, p. 79),—except that for atoms such as H, He, Li this is replaced by a group of two. (*b*) Electrons which are shared may be shared equally or unequally by atoms, so that one can readily account for all intermediate stages between homopolar molecules like H_2 or Cl_2 and polar molecules such as NaCl. (*c*) Besides the cases having a stable group of two, already mentioned, Lewis noted the existence of other probable exceptions to the "rule of eight," namely cases like those of PCl_5, SF_6, $PtCl_6^=$ where a central atom appears to share more than eight electrons with other atoms (reference 13, pp. 102, 114 *et seq.*). Lewis suggested that in such compounds, the shared electrons in excess of eight have "passed into a secondary valence shell" of higher energy.

(2)(*a*) The single bond of the old valence theory of organic chemistry is interpreted as a pair of electrons held jointly in the outer shells of two atoms, and this concept of the chemical bond is extended to inorganic compounds. This electron-pair bond may be symmetrically shared as in H_2 or Cl_2, but more often the electron-pair is nearer one atom than the other. (*b*) The valence of an atom in any molecule is defined as the number of electron pairs which it shares with other atoms (reference 13, p. 104). This is Langmuir's "covalence." Lewis's definition of valence makes the numerical valence the same as the "coordination number" of Werner (reference 13, p. 114). Using this definition, atoms which form a "group of eight" (Langmuir's "octet") by sharing have a valence of four, and "we may regard the maximum valence of four as a sort of norm," although atoms sharing more than four pairs have a higher valence, e.g., S or Pt in SF_6 or $PtCl_6^=$ has a valence of six. When there is an outer shell of eight electrons, there is a strong tendency for these to be all shared. In dealing with ionic or ionized molecules, Lewis uses the term "polar number" (reference 13, pp. 70, 104), e.g., in $Co^{++}Cl^-_2$, the Co atom is said to have a polar number $+2$, the Cl a polar number -1. This is Langmuir's "electrovalence." (*c*) Usually each atomic partner in an electron-pair bond furnishes one of the two electrons, as e.g. in H_2, CH_4, CCl_4, but very often one partner furnishes both: examples, $BR_3 \cdot NH_3$ and other ammonia complexes, where the N atom furnishes both electrons of a pair; $SO_4^=$, ClO_4^-, etc., where the central atom furnishes all the shared electrons except the ionic charges. One-sided sharing as here and in (*a*) causes polarity in the molecule, but does not require a really ionic conception of valence, such as Kossel has used,[14] except in extreme cases. (*d*) Double and triple bonds are considered to involve sharing of two and three pairs of electrons, but are considered to be something very different from merely two or three ordinary electron-pair bonds. Multiple bonds are apparently rarely formed by atoms other than C, N, and O.

The fundamental ideas of Lewis's theory are perhaps the three following: (*A*) each atom (better, each nucleus) in a molecule tends to become sur-

[14] W. Kossel, Ann. d. Physik **49**, 229 (1916); Naturwiss. **7**, 339, 360 (1919); cf. A. E. van Arkel and J. H. de Boer, Chemische Bindung als elektrostatische Erscheinung, S. Hirzel, Leipzig, 1931 (German edition).

rounded by a set of closed shells of electrons; (*B*) shared electrons, forming chemical bonds of the homopolar type, are usually localized between two atoms (or nuclei) which they link together; (*C*) a chemical bond usually consists of a pair of electrons somehow rather closely united to each other. Of these three ideas, Lewis seems to have considered (*C*) as the most fundamental (cf. reference 13, pp. 79–81). Ideas (*A*) and (*B*) were already in use before Lewis's 1916 paper, although Lewis's work greatly developed them and their application. Lewis' most important contribution to the theory consisted perhaps in his introduction of the idea of the completion of closed groups by *sharings*. Idea (*A*), limited, however, to the "group of eight," was proposed by Abegg in 1904 for polar valence and was also used in the same way by Lewis and later by Parson (1915) and Kossel (1916). Idea (*B*) was adapted by Lewis from the concepts of organic chemistry. Langmuir,[15] in further developing the subject in 1919, emphasized idea (*A*) and spoke of the "octet theory" of valence. Bury[16] in 1921 also emphasized idea (*A*).

With the development of the quantum theory, a more fundamental theoretical background has gradually been created for the originally largely empirical ideas of the Lewis theory. Bohr's theory of the periodic system, as modified by Stoner and Main Smith, shows why idea (*A*) is important in valence theory. The group of two found in H^-, He, Li^+ is $1s^2$ and presumably the shared pair of electrons in H_2 is essentially the same, while the highly stable "octet" is ns^2np^6. Bohr's theory also provides for less stable groups of 18 and 32 electrons. Sidgwick[17] showed in a general way in 1923 on the basis of the original Bohr theory how a group of 18, containing *ns*, *np* and *nd* electrons, can be used to interpret complex ions like $PtCl_6^=$ or $Co(NH_3)_6^{+++}$. Revised interpretations which regard the shared electrons in such compounds as belonging to *nd*, $(n+1)s$, and $(n+1)p$, have been given later (cf. Pauling,[18] also I). Knorr[19] has discussed the interpretation of Lewis's theory from the standpoint of the modified Bohr theory.

Bohr's theory was unable to give a satisfactory understanding of the sharing of electrons in molecules, but the new quantum mechanics is showing itself capable of doing this.

London and Heitler, generalizing results obtained from a quantum-theoretical study of the formation of H_2 from H+H, attempted to construct a valence theory which has often been supposed to be the quantum-mechanical equivalent of Lewis's, and which emphasizes Lewis's ideas (*B*) and especially (*C*). This so-called spin theory of valence emphasizes the pairing of electrons and their spins, but deals primarily with the interactions of atoms as wholes. It has, however, not proved very successful.[5,8,10] London made the suggestion that the shared electrons in excess of eight ($ns^2\ np^6$) in

[15] I. Langmuir, J. Am. Chem. Soc. **41,** 868 (1919); Science **53,** 290; **54,** 59 (1921).

[16] C. R. Bury, J. Am. Chem. Soc. **43,** 1602 (1921).

[17] N. V. Sidgwick, Trans. Chem. Soc. **123,** 725 (1923).

[18] L. Pauling, J. Am. Chem. Soc. **53,** 1367, 3225 (1931); **54,** 988 (1932).

[19] C. A. Knorr, Zeits. f. anorg. allgem. Chem. **129,** 109 (1923). Also W. A. Noyes (1917), and others: cf. references given by W. A. Noyes, Zeits. angew. Chem. **44,** 893 (1931).

PCl_5, SF_6 and the like are *nd* electrons; this idea is more specific than the similar idea of Lewis (cf. item 1 (*c*) above), and seems to be a good one.[18,1] One should distinguish carefully[8] between Heitler and London's valence theory and their valuable *perturbation-method for calculating energies* of molecule-formation.

Pauling[18] and Slater[20] have generalized the Heitler and London results for H_2 in a different way, considering more specifically the interactions of pairs of electrons, one electron from each of two atoms. In this way they have obtained important results on valence and valence angles in polyatomic molecules. Their work emphasizes Lewis's idea (*C*) in modified form but focuses attention more on idea (*B*). Their theory seems more limited than Lewis's in that it requires that the two electrons forming an electron-pair bond necessarily come from two different atoms. This makes it inapplicable to many chemical compounds (cf. summary of Lewis's theory, item 2*c*) unless one assumes them to be formed from ions, a procedure which in many cases seems decidedly artificial. The theories of Heitler and London, Pauling and Slater might be called *electron-pairing* theories if Lewis's is called an *electron-pair* theory. It should also be pointed out that the H.L.P.S. electron-pair differs very considerably from Lewis's conception of the electron-pair bond in that the electrons are much less closely associated (see below); in this respect it approaches the truth much more closely than does Lewis's conception.—Pauling and Slater consider a double bond to be merely two ordinary single bonds sticking out from each atom in different directions, and treat the triple bond in a similar way. In this they do not agree very well with Lewis (cf. summary, item 2*d*), nor do they agree with results obtained from the method of molecular orbitals (see below).

Dunkel[21] emphasizes Lewis's ideas (*A*) and (*B*), and describes the shared electrons in molecules in terms of the symbols $[\sigma]$ and $[\pi]$, implying orbital wave functions having properties similar to those of σ and π orbitals in diatomic molecules. The symbol $[\sigma]$, for instance, indicates a molecular orbital concentrated in the region between two nuclei and roughly symmetrical around the line joining them. This is justified by the fact that the shared electron moves in a Hartree field which is roughly symmetrical around this line. Although shared electrons of any given kind nearly always occur in pairs in stable molecules, this seems to be incidental in Dunkel's as in the present work. Unshared electrons are classified as they would be in free atoms. Dunkel's viewpoint evidently resembles that of the present paper. Hund in his recent papers[5,6] has used Dunkel's $[\sigma]$, $[\pi]$ classification, but has used the simple symbols σ, π.

Hund in his important recent papers[5,6,7] concludes that the results obtained by Slater and Pauling using the electron-pair bond method can also be obtained, sometimes more easily, by the method of molecular orbitals, and that the two methods are in many respects equivalent if one restricts the use of molecular orbitals to a type localized between two nuclei,—in agree-

[20] J. C. Slater, Phys. Rev. **37**, 481; **38**, 325, 1109 (1931).

[21] M. Dunkel, Zeits. f. phys. Chem. [**B**] **7**, 81; **10**, 434 (1930).

ment with Lewis's idea (*B*). Reference should be made to Hund's papers for a critical discussion of many points concerning the various quantum theories of valence.

Hund has classified valences according to a number of types, depending in part on whether they are formed by *s* or *p* electrons or involve partial hybridization of *s* and *p* (Hund's *q* type). Hund also classifies chemical bonds according to a number of types, and concludes among other things that single bonds are always of the type $[\sigma]^2$, double bonds of the type $[\sigma]^2\,[\pi]^2$, triple bonds of the type $[\sigma]^2\,[\pi]^4$. It will be noted that this description of double and triple bonds agrees better with Lewis's conclusions (cf. summary, item 2*d*) than does the description given by Slater and Pauling. The absence of free rotation about double bonds can be well accounted for[8,10] by the formulation $[\sigma]^2\,[\pi]^2$, which really goes back to Hückel,[22] while the validity of the explanation given by Pauling and Slater seems doubtful.

Hund shows[6] that the familiar rules of organic chemistry can be explained in terms of the quantum-mechanically predicted behavior of some of the possible types of valence and of bonds. He shows that not all of the theoretically possible types are found in ordinary organic compounds, and shows how this can be understood in terms of energy relations; in particular, by assuming that of the types $[\sigma]$ and $[\pi]$ corresponding to a *p* valence electron, $[\sigma]$ in practise nearly always has the lower energy.

In all this work, Hund's purpose has been to show how the ordinary rules of valence can be derived from the principles of the quantum theory. In doing this, he has found it necessary to specialize the concept of molecular orbitals to the case of *orbitals localized between two nuclei*. Such a localized molecular orbital occupied by two electrons he regards as corresponding to the valence bond of organic chemistry or to that of Lewis, and as being essentially equivalent in the case of ordinary stable molecules to the electron-pair bond of H.L.S.P.

Hund remarks incidentally that the use of localized molecular orbitals is a somewhat poorer approximation than the use of the non-localized orbitals which are advocated in the present series, and illustrates this[6] by discussion of a case which corresponds to CH_2 or H_2O. He also points out that in some cases localized orbitals cannot be used at all, and illustrates this by a consideration of crystal lattices and of the benzene molecule C_6H_6 and related molecules. In his plausible and important interpretation of C_6H_6, Hund uses the same point of view that is emphasized in the present series. His non-localized "π" type in C_6H_6 is very similar to the central-nucleus-viewpoint type $2p[\sigma]$ in $CO_3^=$ or NO_3^- if this type is assumed to be filled by two electrons shared by the O atoms in addition to their ordinarily assumed sharing of one pair each (cf. I, p. 61).

In the present series, idea (*A*) of Lewis's theory is emphasized, idea (*B*) is adopted in generalized form (non-localized molecular orbitals), while idea (*C*) is considered as corresponding more to an incidental than to a really essential characteristic of chemical combination. While the pairing of elec-

[22] E. Hückel, Zeits. f. Physik **60**, 423 (1930).

trons, which is a consequence of the Pauli principle combined with the fact that the electron spin can have two and only two orientations, is very important for the *general theory of atomic and molecular structure*, especially in connection with the formation of closed shells, the importance of the specific role of electron-pairs or of electron-pairing in the *formation of molecules* has, in the writer's opinion, been somewhat over-emphasized, first in Lewis's theory, then in the work of Heitler and London, Pauling and Slater. Reasons for this opinion have been given in a previous paper dealing mainly with diatomic molecules.[8] It seems desirable, however, to take the matter up again in some detail here, approaching it historically.

5. Analysis of the origin and development of the concept of the electron-pair bond

At the outset it is important to notice that Lewis's development of the concept of the electron-pair bond preceded our present knowledge of the electron spin and its properties and of the Pauli principle.

In developing his theory, Lewis noticed that while nearly half of the known atoms contain an odd number of electrons, very nearly all stable molecules contain an even number. This and other information on atomic structure and on molecules suggested that electrons have a tendency to form stable groups of two, held close to each other (cf. reference 13, p. 82 and Fig. 22) by strong forces of unknown character. These forces Lewis for some time considered to be magnetic, since the paramagnetism characteristic of atoms or molecules with an odd number of electrons usually disappears when they combine to give molecules with an even number. Lewis considered that in the formation of a molecule from two atoms containing unpaired electrons, the electrons attract each other in pairs, and that this attraction is, to a considerable extent at least, the cause of molecule formation (cf. e.g., reference 13, p. 149, second paragraph).

The subsequent development of the quantum theory has provided adequate explanation for the above facts concerning electron-pairs. The fact that most of the electrons in atoms as well as in molecules are paired is now well understood in terms of Pauli principle and electron spin. The fact that paramagnetism goes with unpaired electrons, diamagnetism with paired electrons, is also well understood. The fact that electrons are more often unpaired in atoms than in molecules can be understood in terms of the Pauli principle, electron spin, and orbital degeneracy (see below).

According to the quantum theory an electron-pair, in a molecule or in an atom, consists of two electrons, with spins antiparallel, occupying the same orbital and "symmetrically related" so far as their orbital motion is concerned.[8] The magnetic forces between the two electrons are very small while the electrostatic repulsion is large. It is true that the symmetrical relation keeps the electrons nearer together than they would otherwise be, and in this respect shows a similarity to Lewis's concept of a pair of electrons held together by a strong attraction. But, directly contrary to Lewis's idea, the increased nearness of the electrons usually only increases their energy of repulsion.

Thus the formation of a pair of electrons when a molecule is formed, i.e., the occupation of a shared molecular orbital by *two* electrons, far from being *per se* a *cause of combination* of two atoms initially having each an unpaired electron, should tend to prevent combination. According to the present point of view the two electrons concerned act as bonding electrons *in spite of* considerable repulsive forces acting between them. Further details will be found in a discussion of H_2^+ and H_2 in reference 8.

Why then, are unpaired electrons so very much rarer in stable chemical molecules than in atoms? The answer involves several distinct considerations. In the first place, every atom having an odd number of electrons has necessarily at least one unpaired, while as soon as the formation of molecules is permitted, this necessity disappears. This leaves three questions. (1) Why do atoms often have more unpaired electrons than the minimum possible number, which is zero for an even total number of electrons, and one for an odd total number of electrons? (2) Why do stable molecules nearly always have an even number of electrons? (3) Why do such molecules nearly always have all their electrons paired?

The answers to all these questions depend largely on the occurrence and behavior of non-degenerate and of degenerate orbitals. If every orbital in atoms and molecules were decidedly different from every other in energy, then no atom or molecule with an even number of electrons would contain unpaired electrons when in its normal state. The electrons would settle down into the orbitals of lowest energy, with two electrons in each such orbital.

But when there is orbital degeneracy, i.e., when two or more orbitals form a group whose members are equal in energy (cf. section 2 above), then the energy of the electron system as a whole is a minimum if every electron occupies a different orbital belonging to the degenerate group. For example in the nitrogen atom the energy is a minimum when we have the electron configuration $1s^2\ 2s^2\ 2p^3$ and the state ^4S. The orbitals $1s$ and $2s$ are non-degenerate and each is occupied by two electrons, but the type $2p$ has a three-fold degeneracy, being capable for example of giving rise to three distinct orbitals, $2p_{+1}$, $2p_0$, and $2p_{-1}$ in a strong magnetic field. The state ^4S corresponds to one electron, unpaired, in each of the three orbitals $2p_{+1}$, $2p_0$, and $2p_{-1}$. The energy is a minimum for this state because it permits the orbitals of the three electrons to be antisymmetrically related.[8] The resultant spin $S(S=1\frac{1}{2}$ here) is always equal, in molecules as well as in atoms, to $\frac{1}{2}$ times the number of unpaired or antisymmetrically related electrons, here three. If two electrons occupied the same orbital, e.g., $2p_{-1}$, they would necessarily be symmetrically related; but a symmetrical relation, as already noted, forces the electrons together, and would raise the energy above that of the ^4S normal state.

It remains to give reasons why stable molecules nearly always have an even number of electrons, in pairs. In the writers' opinion, this can best be explained in terms of Lewis's idea (*A*) according to which each nucleus in a molecule tends to become surrounded by an atomic closed shell of electrons.

Since such a closed shell, e.g., ns^2 or ns^2np^6, contains an even number of electrons,—for the reason that it is in the last analysis made up of pairs of electrons, although these often belong to degenerate groups, e.g., np^6,—every molecule which satisfies Lewis's idea *A* must automatically contain an even number of electrons, *all paired*. The quantum-mechanical background of idea *A* will be taken up below.

A further reason why molecules with an even number of electrons much more rarely contain unpaired electrons than atoms with an even number of electrons is to be found in the fact that orbital degeneracy is less usual in molecules than in atoms (cf. section 2, above). The O_2 molecule is a good example of an "exception that proves the rule." Here the electron configuration is of a type $\cdots\pi^2$, and the normal state is $^3\Sigma^-$, with $S=1$ and so two unpaired electrons. The π type of molecular orbital has a two-fold degeneracy, so that with two π electrons, one can be π^+, the other π^-, and the two can be antisymmetrically related and so give rise to the normal state with $S=1$.

From the foregoing discussion it will be seen that the empirical evidence which led Lewis to emphasize his idea (*C*), according to which the formation of electron-pairs has a peculiar importance for chemical phenomena, can now all be explained satisfactorily on the basis of general quantum-theoretical considerations without any necessity of adopting idea (*C*).

Nevertheless it must be admitted that most of the striking chemical facts are not inconsistent with an electron-pair bond theory of homopolar valence similar to Lewis's. (His idea of a special attraction between electrons as the cause of pairing must of course be dropped.) A pair of electrons occupying equivalent localized molecular orbitals appears to be on the whole the nearest quantum analogue of Lewis's bond. For the best-known stable molecules the H.L.S.P. electron-pair,—cf. Hund (section 4), and section 13,—is equally good or perhaps better, but for molecules where a pair of electrons is furnished by one atom (cf. Lewis's theory, item 2*c*, and section 6), H.L.S.P. electron-pairs are less suitable than pairs of bonding orbitals. On the other hand for describing molecules having unusually loose binding, or molecules in a high state of vibration or in process of dissociation, provided dissociation would cause the breaking up of electron-pairs into unpaired electrons, electron-pair bonds of the H.L.S.P. type should often constitute a much better approximation than bonding orbitals.

6. Advantages of the present method

The concept of the bonding molecular orbital has the following advantages over the Heitler-London, Slater-Pauling electron-pair bond. (1) It is not necessary to have two electrons to get a chemical bond; one electron occupying a bonding orbital has a bonding effect, although of course not as strong as if two were present.[8] Such "one-electron bonds" are rarely found in stable molecules, it is true, but the concept of bonding orbitals which explains them as a special case can lay claim to greater generality than the electron-pair bond concept which has to be replaced by a different special

concept in order to deal with them. The importance of one-electron bonds[18] is greatly increased if one ventures into the realm of spectroscopically observable molecular entities and of intermediate products in chemical reactions. (2) Bonding molecular orbitals can be constructed for any degree of polarity or unequalness of sharing of electrons between nuclei. H.L.S.P. electron-pair bonds, however, cannot be used for molecules of the extreme polar type, and intermediate cases can be taken care of only by forming linear combinations of wave functions of the polar and electron-pair bond types (cf. section 13). In cases like $Cu(NH_3)_4^{++}$ or CO, the H.L.P.S. method can be used only by starting with assumed states of ionization like $Cu^=$ and NH_3^+ or C^- and O^+, while the method of molecular orbitals can start with Cu^{++} and NH_3 or C and O. For example, the pair of electrons which any N atom in NH_3 shares with the Cu or Pt atom in $Cu(NH_3)_4^{++}$ or $Pt(NH_3)_6^{++++}$ can be assigned to a symmetrically related pair of molecular orbitals which at first belong wholly to the NH_3 but take on more and more of the character of Cu or Pt atomic orbitals as the NH_3 approaches the metal atom. Here the method of molecular orbitals is superior even for loosely bound molecules. These various possibilities make the concept of molecular orbitals more generally useful than that of H.L.S.P. electron-pair bonds. (3) Bonding molecular orbitals are not restricted, like electron-pair bonds, to holding just two nuclei together, but may be distributed between several nuclei. This makes the concept a much more flexible one than that of the electron-pair bond, and makes it possible to account rather directly for a number of phenomena, many of them more or less spectroscopic but others definitely chemical in the ordinary sense, which the electron-pair bond method does not touch. In particular, molecular orbitals can be chosen in conformity to the actual symmetry of the nuclear arrangement, while the electron-pair bond method often disregards (even though it often predicts) this.

In general it may be said that there are many phenomena which can be interpreted in terms of electron-pair bonds only if after setting up these bonds, various *linear combinations* are formed, while the molecular orbital concept goes more directly at the solution, although often seemingly neglecting certain features expressed by the electron-pair concept. It appears probable that in practise one can expect most of the phenomena expressible by the special concept of electron-pair bonds to drop out of the application of the less specialized method using molecular orbitals. This should become clear in the detailed discussion of examples in later papers. It is quite possible, however, that the electron-pair bond method may for many problems be more adapted to quantitative calculations than the present method.

7. Statement and justification of a valence rule

Most of the ordinary numerical aspects of valence, alike for homopolar, heteropolar and for intermediate types of compounds, appear to be expressible by a simple *valence rule* (cf. I)*: Every nucleus in a molecule tends to be surrounded, by means of sharing or transfer of electrons, by an electron distribution corresponding to some stable configuration having a total charge ap-*

proximately equal to or somewhat exceeding the charge of the nucleus. This is essentially idea (*A*) of Lewis's theory together with a generalization of his idea (*B*). This valence rule, it will be noted, is expressed in terms of one-nucleus viewpoints, according to which the state of the electrons near each nucleus is described by means of a set of electron quantum numbers or atomic orbitals associated with the viewpoint of that nucleus.

The important tendency of many atomic nuclei to have all their outer electrons shared (cf. section 4, item 2*b*), as e.g. in the formation of NH_4^+, $BR_3 \cdot NH_3$, or $Cu(NH_3)_4^{++}$ from NH_3 is not accounted for by this valence rule, but its explanation appears to follow incidentally from the application of the present method (cf. discussion of NH_3 and CH_3 in I).

By "stable configuration" in the valence rule is usually meant a set of atomic orbitals completely occupied by electrons (i.e., a set of closed shells) and of such type that further electrons could go only into orbitals of distinctly higher energy. Stable configurations are usually sharply marked off from other configurations with either more or fewer electrons, by virtue of large energy changes that go with a change in principal quantum number, or with the change from a penetrating to a non-penetrating type of orbit even without change in principal quantum number.

Stable configurations which are sharply defined in respect to energy tend to give sharply-defined valence relations. In the transition groups where a *d* shell is in the process of being built up, there is at first no limiting sharply-defined stable configuration, and highly variable valence is the result. As the *d* shell approaches completion, however, the complete group $d^{10}s^2p^6$ apparently begins to serve as a stable limit, which is approached or reached in a number of complex molecules.

A quantum-theoretical justification of the above valence rule follows in part from the fact that, for the electrons in the neighborhood of every nucleus in a molecule, the Pauli exclusion principle makes essentially the same requirements in regard to quantum numbers and closed shells as for the electrons in an isolated atom. This is obviously true for the unshared, inner electrons, and for the ionic structures in definitely heteropolar valence, while its truth for shared electrons appears plausible but requires further investigation. The best evidence for its correctness for shared electrons seems to be the empirical evidence of the success of the rule in interpreting chemical facts. Dunkel[21] has justified his (more or less tacit) use of a similar rule by similar arguments (cf. also Sidgwick, reference 13, p. 98 *et seq.*).

As a result of sharing, the numbers of electrons surrounding certain nuclei sometimes very considerably exceed the numbers in the corresponding neutral atoms. In such cases the sharing is usually relatively loose and presumably is strongly one-sided, although not necessarily ionic. Thus, for example, the electrons shared by the Pt or Cu nucleus in $Pt(NH_3)_6^{++++}$ or especially $Cu(NH_3)_4^{++}$ doubtless belong much more to the N nuclei than to the metal nucleus. The shared molecular orbitals would be approximated by linear combinations containing a relatively small proportion of an orbital,—nd, $(n+1)s$, or $(n+1)p$,—of the metal nucleus and large proportions of $2p\sigma$

orbitals of the N nucleus. Nevertheless the magnetic and other properties of such molecules[18] indicate that even such one-sided sharing is commonly (but not always) strong enough so that, as far as the Pauli principle is concerned, part or all the metal-nucleus-viewpoint orbitals nd, $(n+1)s$, $(n+1)p$ are effectively filled.

In general, the ability of a nucleus to hold more electrons, but a rather definitely limited number more, when in a molecule than in an atom, appears to be conditioned by the fact that shared electrons are, so to speak, electrostatically on a part-time basis for each nucleus while with respect to the Pauli principle they are serving full time.

Proceeding further with the quantum-mechanical justification of the valence rule, one needs next to show why an energy-decrease should occur when atoms or ions so combine that each nucleus becomes surrounded by a "stable configuration." The heteropolar case is fairly well understood. A qualitative understanding also of the energy decrease which occurs in homopolar sharing seems to be obtainable in terms of one-nucleus viewpoints.

8. Cause of homopolar valence forces

London and Heitler concluded that the previously mysterious homopolar valence forces are explained by "resonance" or "exchange" integrals,—like $\mp I_2$ in Eqs. (2), (3) below and similar integrals occurring in the electron-pair case,—which give rise to sharp attractions or repulsions. As will be seen in section 10, these energy integrals correspond to the fact that bonding molecular orbitals give a higher, anti-bonding orbitals a lower, electron density in the regions between nuclei than if they were formed by mere overlapping of the electron densities of the atomic orbitals from which they might be formed. In a certain sense the above simple explanation of the homopolar valence forces is an adequate one. But it is of interest to see if one can go farther in understanding the matter physically by seeking to find reasons *why* these changes in electron density and in energy should occur when molecular orbitals are formed.

The simple case of H_2^+ is instructive.[8] If we let $E.E.$ represent the energy of the electron, taken as zero for $R=\infty$, and $N.E.$ the energy of repulsion of the nuclei, then the energy change as the two H nuclei come together is

$$\Delta E = N.E. + E.E.$$

If $E.E.$ decreases with decreasing R considerably faster than $N.E.$ increases, until R reaches a fairly small equilibrium value, we have a stable molecule (negative ΔE). If it increases, or decreases less rapidly than $N.E.$ increases, we have repulsion. The two states $\sigma 1s$ and $\sigma^* 1s$ of H_2^+ correspond to these two cases. The exchange energies in the two cases are proportional respectively to $-I_2$ and $+I_2$.

The fact that ΔE is negative for $\sigma 1s$ and positive for $\sigma^* 1s$ corresponds to a rapid decrease of $E.E.$ with R for $\sigma 1s$ and a very much less rapid decrease for $\sigma^* 1s$. A good physical reason for this can be seen when we note that as $R \rightarrow 0$, the $1s$ atomic orbital of $H+H^+$ must shrink through $\sigma 1s$ to a much

more concentrated $1s$ orbital of the united-atom, while through σ^*1s one reaches $2p\sigma$ of the united-atom, which is of about the same degree of concentration and has the same energy as $1s$ of $H+H^+$.

These considerations suggest that the behavior of bonding molecular orbitals like $\sigma 1s$ of H_2^+ is, qualitatively at least, essentially the result of an increase in effective nuclear charge without change of quantum numbers (no promotion), as R decreases. This effect must of course be intense enough to make ΔE negative in the above equation over a considerable range of R values. The fact that it actually does so for $\sigma 1s$ of H_2^+ could hardly have been predicted from our qualitative explanation, but this does not render the latter invalid, although it does show the superiority of the Heitler-London method for quantitative prediction. With anti-bonding orbitals like σ^*1s, the changed quantum numbers (promotion) more or less neutralize the effect of the increasing effective nuclear charge, thus permitting $N.E.$ always to exceed the negative of $E.E.$ in the above equation.

A somewhat more complicated case than H_2^+ is that of the N_2^+ molecule (cf. also section 10). The formation of the molecule in its normal state may be expressed as follows:

$$N(1s^2 2s^2 2p^3) + N^+(1s^2 2s^2 2p^2) \rightarrow N_2^+(1s^2 1s^2 \sigma 2s^2 \sigma^* 2s^2 \pi 2p^4 \sigma 2p).$$

If we let $R=\infty \rightarrow R=R_e \rightarrow R=0$ $(N+N^+ \rightarrow N_2^+ \rightarrow Si^+)$, we should probably have[9]

$$N[2s] \rightarrow N_2^+[\sigma 2s] \rightarrow Si^+[2s],\ N[2s] \rightarrow N_2^+[\sigma^* 2s] \rightarrow Si^+[3p(\sigma)],$$
$$N[2p(\sigma)] \rightarrow N_2^+[\sigma 2p] \rightarrow Si^+[3s],\ N[2p(\pi)] \rightarrow N_2^+[\pi 2p] \rightarrow Si^+[2p(\pi)].$$

Here the bonding molecular orbitals $\sigma 2s$ and $\pi 2p$ behave like $\sigma 1s$ of H_2^+ in that they are unpromoted as $R \rightarrow 0$, while the anti-bonding orbital behaves like σ^*1s of H_2^+ in that it is promoted as $R \rightarrow 0$. But $\sigma 2p$, which is promoted as $R \rightarrow 0$, is a bonding orbital, like $\pi 2p$ which is unpromoted. These results indicate that the difference between bonding and anti-bonding orbitals is not simply a difference between an unpromoted and a promoted condition if promotion is defined according to what happens in the united-atom ($R=0$).

If, however, we *redefine promotion* in terms of the *one-nucleus viewpoints* of the nuclei in the molecule, introducing the new word "*premotion*" in order to avoid misunderstanding, *bonding electrons* are essentially *unpremoted electrons* or (from an energy standpoint) sometimes slightly premoted electrons, while *anti-bonding electrons* are (strongly) *premoted electrons*. A few examples will make clear the exact sense in which premotion is here defined. According to the one-nucleus viewpoint of either of the two N nuclei in N_2^+ (or N_2), the seven (or eight) electrons occupying the bonding molecular orbitals $\sigma 2s$, $\sigma 2p$, and $\pi 2p$ are reckoned as $2s$, $2p\sigma$, and $2p\pi$ electrons, hence all unpremoted; in N_2 they form, from the point of view of each nucleus, a complete, although completely shared, L shell. The anti-bonding σ^*2s molecular electrons, however, must be reckoned as 3-quantum, hence premoted, electrons. In HCl, two bonding electrons may be considered as shared by the

H and Cl nuclei. From the H nucleus viewpoint, these represent $1s^2$, from the viewpoint of the Cl nucleus they are $3p\sigma^2$. But from both viewpoints they are unpremoted in the sense that their quantum numbers are the same as those of electrons already present in the H and Cl atoms. In $PtCl_6^=$ (cf. I), two bonding electrons may be considered as shared between each Cl and the Pt nucleus. From the point of view of each Cl nucleus, the electrons it shares are $3p[\sigma]^2$. From the point of view of the Pt nucleus, the twelve electrons it shares are $5d_\gamma{}^4 6s^2\ 6p^6$, while $5d_\epsilon{}^6$ is also present but unshared. The shared $5d$ and $6s$ electrons are unpremoted, the $6p$ electrons are premoted in the sense that the Pt atom alone, in its normal state, has a configuration $5d^9 6s$ without $6p$ electrons. But the fact that $6p$ electrons are only a little less easily ionized than $5d$ or $6s$ in the Pt atom makes it possible for the Pt nucleus in $PtCl_6^=$ to use them as bonding electrons of a slightly premoted type.

The results of the preceding paragraphs can be formulated in part as follows. *Chemical combination of the homopolar type is a result of the shrinkage and consequent energy-decrease of atomic orbitals in the fields of neighboring nuclei, when such orbitals are shared with little or no premotion.*

A condition which is necessary for successful sharing is that the binding energy shall not be too different in the atomic orbitals involved. Otherwise the sharing becomes a one-sided affair: either an electron is almost completely transferred from one nucleus to another, or else almost no sharing or transfer at all takes place (cf. Hund[5] for further discussion and details). Polar molecules like Cs^+F^- approximate the former and loosely bound molecules like $Cu(NH_3)_4^{++}$ the latter case; unstable molecules like HeH are extreme examples of the latter case.

9. Valence saturation

The important phenomenon of valence saturation can now be somewhat understood. It should first be pointed out that saturation is usually a relative rather than an absolute matter. Whether an atom or molecule acts in a saturated manner usually depends on the other atoms or molecules with which it is placed in contact, and furthermore it depends very much on circumstances such as temperature, pressure, and presence or absence of light. In practise, however, a molecule is generally considered saturated if it is stable at room temperature in the presence of others of its own kind and of air and diffused sunlight and perhaps of certain common chemicals. With this definition, a given atom can occur in many different saturated molecules.

A state of saturation usually exists when each nucleus in a molecule is surrounded by a set of closed shells which is stable with respect both to gain and to loss of electrons, including gain by sharing. A set of closed shells is stable toward loss of electrons if it has a high ionization potential; toward gain of electrons if a fairly large energy would be required to transfer an electron to an orbital outside the set. Closed shells ending with ns^2np^6, also the shell $1s^2$, usually fulfill both criteria, while closed shells ending with nd^{10}

fulfill neither since the ionization potential is relatively low and the energy of excitation required to get nd^9 $(n+1)s$ is always small, or even negative. Even "stable" closed shells do, however, often take part in further sharing, usually by allowing two of their electrons to be shared by another nucleus (cf. section 4, item 2*b* and the examples of NH_3 combinations cited in section 7). Closed shells ending with ns^2, $n>1$, are relatively stable toward gain of electrons, but not toward loss. Atoms which are surrounded by stable closed shells when alone do not form molecules, because their electrons are too strongly held for either transfer or sharing and at the same time their outer shells contain no vacant places of low energy for receiving transferred or shared electrons. From the preceding discussion it will be seen that the property of saturation is not a property of closed shells as such, but depends on the occurrence of marked energy discontinuities, which often but not always are associated with closed shells.

Electrons are not accepted beyond the point of saturation simply because the unfavorable energy effect corresponding to premotion to a vacant position is greater than the normal favorable energy effect corresponding to sharing without premotion. The quantum-mechanical forces which cause saturated molecules to repel other molecules if they come too close may be described as *premotion forces*. The action of the premotion forces in polyatomic molecules is of the same nature as the anti-bonding action [8,9] of anti-bonding electrons in diatomic molecules.

In diatomic molecules there is sometimes a forced sharing of anti-bonding electrons, incidental to the sharing of bonding electrons, which may be considered as a sort of supersaturation. That is, electrons are included in the molecule in promoted orbits which would not be expected from the valence rule. An example is the N_2 molecule with the electron-configuration $1s^2$ $1s^2$ $\sigma 2s^2$ $\sigma^* 2s^2$ $\pi 2p^4$ $\sigma 2p^2$. From the point of view of each nucleus, that nucleus is surrounded by $1s^2$ unshared, then by $2s^2$ $2p\pi^4$ $2p\sigma^2$ (shared, cf. section 8), forming a complete K and L shell, and in addition by two (partially) premoted electrons probably classifiable as 3-quantum electrons, corresponding to the shared $\sigma^* 2s^2$. The molecule NO is even more supersaturated, by the inclusion of a $\pi^* 2p$ electron in its configuration.[9] This functions very decidedly as a premoted 3-quantum electron from the point of view of one or both nuclei.

More commonly one finds an opposite condition in which nuclei in molecules are surrounded by less than a complete stable shell. This often results from a competition of different nuclei for electrons, or an inability of certain nuclei to share enough electrons to complete the shells of all the other nuclei. Probable or possible examples of nuclei with incomplete outer shells are B in BCl_3, C in $CO_3^=$, N in NO_3^-, S in SO_3 (cf. I).

Comparison of Properties of Molecular Orbitals and Electron-pair Bonds

10. Properties of molecular orbitals

It has been stated in previous sections that the concept of molecular

orbitals possesses all the important properties which make the concept of electron-pair bonds valuable in explaining and predicting molecular structures, and further, is less specialized and so adapted to explain a wider variety of phenomena. The characteristics of the two concepts can be well understood by developing them for the simple molecules H_2^+ and H_2, then generalizing.—The reader should also refer to Hund's valuable discussion[5,6] of molecular orbitals (Hund's concept *c* in reference 5) and their relation to electron-pair bonds (Hund's concept *b*).

Pauling[23] was the first to apply the Heitler-London method, developed originally for H_2 and there yielding among other things the Heitler-London electron-pair bond, to the one-electron molecule H_2^+. Applied to H_2^+, it helps one to understand molecular orbitals. One finds that when $H(1s)$ and H^+ approach, either attraction with formation of normal H_2^+, or repulsion, may occur. The two modes of interaction correspond to the two energetically distinct states one would have for $H(1s)+H^+$ if the two H nuclei were slightly different in charge.

Let the two H nuclei be designated A and B and the corresponding $1s$ atomic orbitals ϕ_A and ϕ_B. The molecular orbitals resulting when H and H^+ are brought together can be expressed in zeroth approximation as follows:

$$\phi_0 = c_0(\phi_A + \phi_B), \qquad \phi_1 = c_1(\phi_A - \phi_B), \tag{1}$$

where the normalizing factors c depend on the distance R between the nuclei. The energy changes ΔE are in first approximation as follows:[23]

$$\phi_0 : \Delta E = e^2/R + [e^2/(1+S)](-I_1 - I_2) \tag{2}$$

$$\phi_1 : \Delta E = e^2/R + [e^2/(1-S)](-I_1 + I_2). \tag{3}$$

Here

$$I_1 = \int [\phi_B^2/r_A]d\tau, \qquad S = \int \phi_A\phi_B d\tau, \qquad I_2 = \int [\phi_A\phi_B/r_A]d\tau, \tag{4}$$

where r_A is the distance of the electron from nucleus A. The "resonance" term $\mp I_2$ turns the scales in favor of attraction in the case of ϕ_0, of repulsion in that of ϕ_1.

The orbital ϕ_0, whose presence leads to formation of a stable molecule, belongs to the class of bonding orbitals and may in fact be taken as the prototype of these, at least for diatomic molecules with equal nuclei. Similarly ϕ_1 may be taken as a prototype for diatomic anti-bonding orbitals. It will be noted that ϕ_0 is symmetrical in the nuclei while ϕ_1 is antisymmetrical, and that ϕ_0 is nodeless while ϕ_1 has a nodal plane half-way between the nuclei.

The behavior of the ΔE's for ϕ_0 and ϕ_1 (Eqs. 2, 3) is easily understood when one considers the mean charge density $e^2\phi^2$ corresponding to each. One finds

$$\phi_0^2 = c_0^2(\phi_A^2 + \phi_B^2 + 2\phi_A\phi_B), \qquad \phi_1^2 = c_1^2(\phi_A^2 + \phi_B^2 - 2\phi_A\phi_B). \tag{5}$$

[23] L. Pauling, Chem. Rev. **5**, 173 (1928).

This shows that $e^2\phi_0^2$ is relatively more concentrated in the region between the nuclei than if one had a mere overlapping of the densities $e^2\phi_A^2$ and $e^2\phi_B^2$, which would give $\frac{1}{2}e^2(\phi_A^2+\phi_B^2)$ for $e^2\phi^2$. On the other hand ϕ_1 is much less concentrated between the nuclei than for a mere overlapping of $e^2\phi_A^2$ and $e^2\phi_B^2$. With mere overlapping of densities, one would have a slight repulsion between H and H^+, corresponding to $e^2/R-e^2I_1$ in Eqs. (2) and (3). It is the added concentration between the nuclei or withdrawal from this region, in ϕ_0 and ϕ_1, which expresses itself in $\mp I_2$ in Eqs. (2) and (3) and is decisive in making ϕ_0 bonding and ϕ_1 anti-bonding.

The above results for molecular orbitals in H_2^+ can be generalized to other molecules. N_2^+, obtained by the union of $N^+(1s^2\ 2s^2\ 2p^2)+N(1s^2\ 2s^2\ 2p^3)$, will serve as an example. Strictly, we should use also a certain admixture of $N^{++}+N^-$, and should consider various excited states of N^++N, but it is unlikely that this would seriously affect the results. Also, it should be noted that in this discussion we are treating the group of three states of N^+ (^3P, ^1D, and ^1S) and of $N(^4S, ^2D, ^2P)$, which arise from the electron configurations mentioned, as if each group were a single state. This is, however, legitimate because of the large energy of formation of N_2^+. N_2^+ in its normal state contains one $\sigma 2p$ electron, which may be considered, approximately at least, as going over on dissociation to a $2p$ atomic orbital of the neutral N atom. This may be on either nucleus A or nucleus B, so one gets approximately

$$\phi_{\sigma 2p} = c(\phi_{A2p} + \phi_{B2p}),$$

where of course it is understood that the $2p$ atomic orbitals are based on a suitable atomic Hartree field. Instead of the bonding orbital $\sigma 2p$, one could also get the anti-bonding orbital σ^*2p, which should occur in some excited, perhaps unstable, state of N_2^+: $\phi_{\sigma^*2p}=c^*(\phi_{A2p}-\phi_{B2p})$. The bonding orbitals $\sigma 2s$ and $\pi 2p$ in N_2^+ could similarly be shown to be capable of being respectively approximated by sums $\phi_A+\phi_B$ made up of atomic $2s$ or $2p$ orbitals, and the anti-bonding orbital σ^*2s could be approximated by a difference $\phi_A-\phi_B$.

Also in *unsymmetrical diatomic molecules*, molecular orbitals can be approximated by linear combinations of atomic orbitals: bonding type, $\phi=a\phi_A+b\phi_B$. In strongly ionic molecules, $a\gg b$ or vice versa. The results stated in this and the preceding paragraph have already been given by Hund[5,6] and need not be further discussed here.

It seems clear that molecular orbitals in polyatomic molecules also can be approximated as linear combinations of atomic orbitals (cf. I for examples), and that bonding orbitals are additive linear combinations. In general, polyatomic bonding orbitals may be expected to surround several nuclei, although in special cases such orbitals may be largely confined to the neighborhood of two nuclei.

11. Maximum overlapping as a criterion of bonding power

It is now of interest to learn how the bonding power of a molecular orbital depends on its form. This can be seen from expressions like that for ϕ_0 in

Eq. (1). Since the lowered energy of bonding orbitals results from the fact that they are relatively concentrated between the nuclei, one readily sees that in order to have the bonding energy as large as possible, products such as $\phi_A\phi_B$ in Eq. (4) should be as large as possible, which is true if the atomic orbitals ϕ_A and ϕ_B overlap as much as possible. This can also be seen directly from the energy equations (cf. Eqs. 2, 4). This is Pauling's and Slater's *criterion of maximum overlapping of atomic orbitals*. The process of applying this criterion to the problem of determining the zeroth approximation atomic orbitals which give *strongest bonding* (Pauling's "best bonding eigenfunctions"), when ϕ_A and (or) ϕ_B belong originally to degenerate orbitals, seems to be the essence of the method used by Slater and Pauling (cf. reference 20, p. 1141).

Slater's and Pauling's original derivations of this criterion from the equations for electron-pair bonds (cf. Eq. (8) and related ΔE equations) give the impression that the existence of the criterion depends on the overlapping of the *orbits* of *two actual electrons*. The present approach shows, however, that the criterion for finding a *best bonding molecular orbital* for a pair of atoms depends only on the overlapping of two atomic orbitals, regardless of whether the resulting molecular orbital is occupied by two electrons, by one electron, or by none; the validity of the procedure, used in the present series, of regarding such an orbital occupied by two electrons as essentially equivalent to an electron-pair bond will be considered shortly.

Furthermore, the present approach shows that the criterion of maximum overlapping is just as applicable to molecular orbitals which are approximate linear combinations of atomic orbitals of several atoms as to those which are formed from orbitals of just two atoms. The present use of molecular orbitals which connect several nuclei is perhaps the most essential difference between the present method and that of Slater and Pauling.

12. Doubly- or multiply-occupied molecular orbitals

In nearly all chemically stable molecules, every bonding molecular orbital which is occupied at all is occupied by the full quota of electrons allowed by the Pauli principle, namely two if the orbital is non-degenerate, $2n$ if it is degenerate and n is the degree of degeneracy. The most important properties of such multiply-occupied bonding orbitals can easily be generalized from a consideration of the doubly-occupied $\sigma 1s$ orbital in normal H_2. Omitting constant factors and spins, the wave function χ for normal H_2 may be approximated by the following expression:

$$\chi = \phi_0(1)\phi_0(2) \tag{6}$$

Eq. (6) implies that the two electrons move entirely independently, in the $\sigma 1s$ orbital, here denoted by ϕ_0; the numbers 1 and 2 refer to the coördinates of the two electrons. This, however, is true only as a rough approximation. Although the agreement of Eq. (6) with the truth can be made rather good by making the best possible choice of the form of ϕ_0,—the simple form given in Eq. (1) can be greatly improved on by adjusting it to correspond

to a one-quantum orbit in a suitable Hartree field,—it can never be made exactly right without adding a correction term. This is necessary in order to take care of detailed effects of the two electrons on each other's motions. The same kind of problem is met with in atomic spectra, for example in the He atom. Eq. (5) is a fairly good approximation for this atom if ϕ_0 is taken to refer to the 1*s* atomic orbital in a suitable Hartree field.

Slater[24] has found, in a paper preceding the work already referred to above, that if ϕ_0 is approximated by $c_0(\phi_A+\phi_B)$ in accordance with Eq. (1), the corrected expression for χ can be closely approximated by:

$$\chi = a\phi_0(1)\phi_0(2) - b\phi_1(1)\phi_1(2), \tag{7}$$

where ϕ_1 is as given in Eq. (1). For the molecule in its equilibrium condition ($R=R_e$), Slater finds $a/b=8$, approximately; but for a molecule in the process of dissociation ($R\to\infty$), a/b approaches 1. This indicates that Eq. (6), even with ϕ_0 taken as $c_0(\phi_A+\phi_B)$, and of course all the more so if a better ϕ_0 is used, is a good approximation for H_2 in its ordinary stable state.

Reflection indicates that it is safe to generalize the type of approximation contained in Eq. (6) to unsymmetrical and to polyatomic molecules (cf. also reference 5). It even appears that the approximation may often be relatively better for polyatomic molecules than for H_2. Hence the general use of molecular orbitals in describing shared electrons in stable molecules seems to be justified as a good approximation. In the case of chemical bonds with large R and small dissociation energy, however, cases may occur where the use of molecular orbitals as in Eq. (6), although still formally possible, does not constitute a good approximation.

The nature of the correction required in Eq. (6) is somewhat different for the case of H_2 than for that of He, because of the non-centralness of the Hartree field in H_2. The nature of this difference can be seen if we consider what would happen if we increased R to a large value in H_2. One would then find that the molecule would usually be very much like two H atoms with one electron on each, which means that χ would be essentially of the form $\phi_A(1)\phi_B(2)$ or $\phi_A(2)\phi_B(1)$. There would, however, be a small possibility of finding both electrons attached to one H nucleus, in which case χ would be essentially $\phi_A(1)\phi_A(2)$ or $\phi_B(1)\phi_B(2)$, corresponding to H^-+H^+ or $H^+ +H^-$. Returning to $R=R_e$, it would seem reasonable that the true χ should be approximately a linear combination of the four expressions just given.

13. Comparison with Heitler-London electron-pair bond

Heitler and London's original approximation[25] for normal H_2, omitting constant factors, was

$$\chi = \phi_A(1)\phi_B(2) + \phi_A(2)\phi_B(1). \tag{8}$$

[24] J. C. Slater, Phys. Rev. **35**, 514–5 (1930).

[25] W. Heitler and F. London, Zeits. f. Physik **44**, 455 (1927); Y. Sugiura, Zeits. f. Physik **45**, 484 (1927).

This expresses a state of affairs in which if one electron is in ϕ_A, the other is necessarily in ϕ_B. For large R, this would keep the electrons separated, with one near each nucleus, but for $R = R_e$, where ϕ_A and ϕ_B overlap strongly, the effect is much less pronounced.

Slater[24] has found that the true state of H_2 for $R = R_0$ is better expressed by

$$\chi = \alpha[\phi_A(1)\phi_B(2) + \phi_A(2)\phi_B(1)] + \beta[\phi_A(1)\phi_A(2) + \phi_B(1)\phi_B(2)], \quad (9)$$

with α/β roughly equal to 8. [That the ratio α/β here is nearly the same as a/b in Eq. (7) is accidental.] Eq. (9) allows the two electrons to be close together near the same nucleus oftener than Eq. (8) would permit.

Eq. (9) is a step from Eq. (8) toward Eq. (6); Eq. (6) implies that the probability of any stated position for one electron is independent of the position of the other. On the other hand, Eq. (7) is a step from Eq. (6) toward Eq. (8), if ϕ_0 and ϕ_1 are as given in Eq. (1). In fact, Eqs. (7) and (9) become identical if Eq. (1) is used and if the coefficients a, b, α, β, are properly related. Slater (private communication concerning details of Ref. 25) found that for $R = R_e$ in H_2, the ratio of the coefficients c_0 and c_1 in Eq. (1) is about that given by $(c_1/c_0)^2 = 6.21$. Substituting this, also $a/b = 8$, in Eq. (7), the latter reduces to the form of Eq. (9) with $\alpha/\beta = 14.21/1.79$, which happens to be very nearly the same as a/b.

Eq. (8) is the expression, in terms of wave functions, of the Heitler-London, Pauling-Slater concept of the electron-pair bond, for the case of H_2. Eq. (6) is the expression in terms of wave functions, for H_2, of the concept of a bonding orbital occupied by two electrons. Slater's Eqs. (7) and (9), in view of his result that a/b and α/β are nearly equal, suggest that the two concepts, while departing from the truth in opposite ways, are about equally good approximations to it in their ability to describe the positions of two equivalent bonding electrons relative to each other and to two nuclei. Other considerations, pro and con, may enter besides the values of the ratios a/b and α/β, and of course such ratios would be different for other examples than H_2, but it seems fairly safe to assume that the molecular orbital concept ordinarily does not give too bad an approximation in its description of the electron positions. We have already seen that in two other important properties,—the fact of concentration between the nuclei and the existence of the criterion of maximum overlapping,—the concept of bonding molecular orbitals gives the same results as that of electron-pair bonds.

For electron-pair bonds, the first of these properties follows if one determines the electron density for a single electron; this is given by $\int \chi^2 d\tau_2$ and comes out proportional to

$$\phi_A^2 + \phi_B^2 + 2S\phi_A\phi_B, \quad (10)$$

where S is as in Eq. (4). The concentration term here is less by a factor S than in Eq. (5). When one takes $\int \chi^2 d\tau_2$ using Eq. (6), however, one gets the same result as is expressed in Eq. (5). The existence of the criterion of maximum overlapping follows in the case of the electron-pair bond by a considera-

tion of the above electron density expression, or of ΔE expressions,[25] in much the same way that it does for bonding orbitals.

It is instructive to compare the application of the method of H.L.P.S. electron-pairs and that of molecular orbitals to molecules which have unequal nuclei and are partly polar, such as LiH, HF or HCl. If one uses electron-pair bonds, one forms a linear combination of polar and electron-pair wave functions:[18]

$$\chi = \alpha[\phi_A(1)\phi_B(2) + \phi_A(2)\phi_B(1)] + \beta\phi_B(1)\phi_B(2), \qquad (11)$$

where B is the more negative atom. Using molecular orbitals, Eq. (6) applies, with

$$\phi_0 = a\phi_A + b\phi_B, \; b > a. \qquad (12)$$

Using Eq. (12), Eq. (6) becomes identical with Eq. (11) if we put $ab = \alpha$, $b^2 = \beta$, except that Eq. (6) gives an extra term $a^2\phi_A(1)\phi_A(2)$. This term is relatively unimportant if the molecule is strongly polar; for example if $b = 2a$, the coefficient of $\phi_A(1)\phi_A(2)$ is only one fourth that of $\phi_B(1)\phi_B(2)$. As in the case of equal nuclei discussed above, however, the truth must lie between Eq. (11) and Eqs. (6), (12). If both forms represent about equally good approximations, as seems likely, Eq. (6) may well be preferred for many purposes because of its simpler conception and formulation.

Evidence of Space Quantization of Atoms upon Impact

By H. Kuhn and O. Oldenberg
Zweites Physikalisches Institut, Goettingen, and Physical Research Laboratory, Harvard University

(Received May 16, 1932)

The present paper offers an interpretation of certain diffuse bands observed in the spectra of metal vapors mixed with rare gases. Two separate maxima on the short wave-length side of the resonance line of the metal are not due to vibrational quanta as in the case of other similar diffuse spectra. Instead they are interpreted as separate electronic levels of the pair of colliding atoms described by space quantization taking place in the mutual approach. The diffuse structure on the long wave-length side is due to vibrational quanta of molecules.

I. Experimental Facts

SEVERAL years ago, some fluorescence experiments were performed,[1] the aim of which was to investigate the interaction in a single elementary process of the three kinds of energy: light, excitation energy, and kinetic energy of atoms. It was expected that an excited atom, when colliding with a rare gas atom, might react in two ways: either it might *split up* its energy of excitation into two parts, radiation and kinetic energy of the two atoms, thus giving rise to a continuous spectrum spreading from the resonance line to the long wave-length side; or instead the kinetic energy of impact might *join* with the excitation energy thus producing radiation that belongs to a continuum on the short wave-length side of the resonance line.[2]

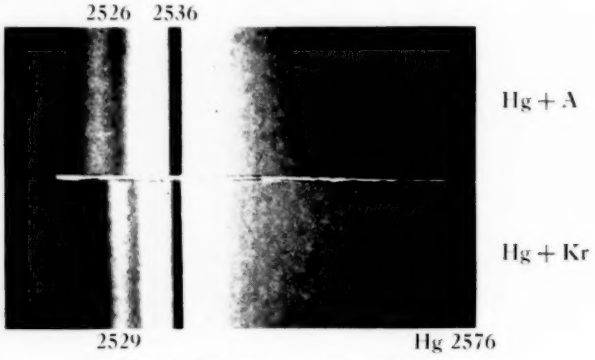

Fig. 1. Mercury—rare gas bands.

In order to check this type of interaction by experiment, the fluorescence radiation of mercury vapor with a large amount of a rare gas added has been investigated. In this experiment, the excited mercury atom has a good chance

[1] O. Oldenberg, Zeits. f. Physik **47,** 184 (1928); **51,** 605 (1928); **55,** 1 (1929).

[2] A theoretical discussion of this process has recently been given by V. Weisskopf, Zeits. f. Physik **75,** 287 (1932) and H. Margenau, Phys. Rev. **40,** 387 (1932).

to collide during its lifetime with a rare gas atom. Actually, continuous spectra have been observed on both sides of the resonance line of mercury. Under most conditions, however, instead of having a uniform distribution of intensity, they showed an unexpected diffuse structure as seen in Fig. 1. (The intense Hg line 2537 itself was taken out of the spectrum by an absorption tube with Hg vapor, in order to prevent scattering of light at the photographic plate.) When A or Kr is added to Hg vapor, two broad diffuse maxima appear at the short wave-length side of the resonance line and a finer diffuse structure on the long wave-length side. Only a trace of structure has been observed in Hg+Ne and no structure in Hg+He or Hg+Xe.

The same type of structure has recently been observed by Krefft and Rompe[3] in mixtures of thallium and other metalic vapors with rare gases. It is of interest that also with He added they obtained two well-defined separate maxima, even broader than with the other rare gases.

II. Criticism of the Interpretation as a Band Spectrum

The fine structure on the long wave-length side has been explained by the formation of molecules from mercury and rare gas atoms. In the fluorescence experiments only the mercury, not the rare gases, was excited. It must be assumed that the excited mercury atom attracts the rare gas atom by the polarization force (cf. section IV). The conspicuous broad maxima on the short wave-length side, however, which present quite a different aspect, have not been explained satisfactorily as yet. In order to interpret them in the usual

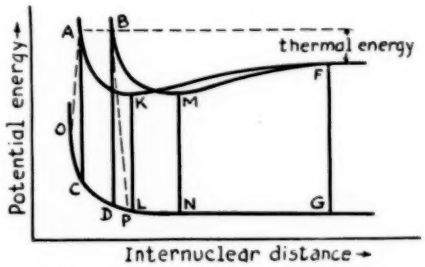

Fig. 2. Hypothetical band structure. *AC*, molecular band; *FG*, atomic line.

way as vibrational quanta of a band, potential curves must be assumed as in Fig. 2. The short wave-length position of the maxima must be represented by a longer vertical distance *AC* than *FG* representing the line emission. It must be assumed that excited molecules with strong vibration combine with normal molecules with no vibration or with just one quantum. It is not evident, however, why the vibration of the *excited* molecules is restricted to such a small range of high energy near *A* that the vibrational structure of the *normal* state comes out clearly.

It has been suggested[4] that the diagram of Fig. 2 does not apply because it fails to represent the *rotational* energy of the molecule, which can exceed

[3] H. Krefft and R. Rompe, Zeits. f. Physik **73**, 681 (1932).

[4] Oldenberg, Zeits. f. Physik **55**, 10 (1929).

considerably the dissociation energy. This argument cannot be maintained because of the selection rule. In the emission process the angular momentum of the molecule changes only by the fundamental unit $h/2\pi$. For heavy molecules as HgA or HgKr this corresponds to such a small change of rotational energy that the single band cannot spread far from its zero line. Hence the rotational energy will not affect the position of the bands to any considerable extent.

These criticisms based on a rather vague argument are strongly supported by the new observations of Krefft and Rompe, who observed the same type of structure with a particularly broad extension in the mixture of Tl vapor and He. The frequency difference between the two maxima is about 500 cm^{-1}, that is of the order of magnitude of the vibrational quantum of a regular molecule.[5] Now it cannot be assumed that the molecule TlHe exists with a considerable energy of dissociation (to be observed as a large quantum of vibration) and that it plays a part in the discharge in 800°C, for the He atom has by far the smallest polarizability[6] of all rare gases (by far the lowest condensation point, the smallest heat of vaporization, the least adsorption). Therefore the type of structure as observed in Tl+He on the short wave-length side of the Tl line must be interpreted by an entirely different process.

III. Radiation during Collision; Space Quantization

It seems that this process can be understood on the basis of the theory of band spectra without an arbitrary hypothesis. It must be assumed that the excess energy of the short wave-length diffuse structure, not explained as dissociation energy of molecules, is due to the kinetic energy of thermal impacts.[7] The mercury atom emits this radiation instead of the sharp resonance line at the moment of a collision with a rare gas atom. This is just the process mentioned above, for the study of which the experimental work was undertaken. This process has been described on the basis of potential curves by Winans and Jablonski.[8] By the same mechanism Weizel[9] recently explained the asymmetric short wave-length broadening of some helium lines discovered by Hopfield.[10] Weizel has succeeded in connecting this broadening with the

[5] The correction discussed by H. Kuhn (Zeits. f. Physik **63,** 458 (1930) and **72,** 462 (1931)) cannot change the order of magnitude of the observed difference, while still discrete maxima of intensity are observed.

[6] J. C. Slater and J. G. Kirkwood, Phys. Rev. **37,** 682 (1931); H. Margenau, Phys. Rev. **37,** 1425 (1931).

[7] From this assumption it follows that Condon's wave-mechanical interpretation of diffuse maxima (Phys. Rev. **32,** 866 (1928)) does not apply to the spectrum under discussion. Condon considered a sharply defined initial state (lowest vibrational level of the upper electronic state) and a final state which belongs to a continuous range. He concluded that diffuse maxima might appear closely connected with the waves representing the motion in the final state. This theory does not apply to the conditions described by Fig. 2, because here the initial state, too, is not sharply defined but is itself diffuse due to the velocity distribution. Moreover the obseved maxima seem to be broader than those described by Condon's theory.

[8] T. G. Winans, Phil. Mag. **7,** 558 (1929). A. Jablonski, Zeits. f. Physik **70,** 723 (1931).

[9] W. Weizel, Phys. Rev. **38,** 642 (1931).

[10] J. J. Hopfield, Astrophys. J. **72,** 133 (1930).

molecular spectrum. Discussing the broadening of the He line 585 on the basis of the theory of band spectra, he finds one potential curve that describes the interaction of two normal He atoms colliding with each other, and four separate curves for an excited 2^1P atom colliding with a normal one. Due to a selection rule only two of these four curves, both representing repulsion, combine with the normal state (Fig. 3).

It can be assumed that in the normal state the Hg atom has a smaller effective cross section than in the excited state. Hence in the diagram the potential curve of the normal state remains flat even for a small internuclear distance. Thus the kinetic energy of the collision—at the very moment in which it is turned into potential energy (A in Fig. 3)—takes part in radiating a larger quantum AC than the excited atom would radiate by itself (FG). The superimposed effect of the two potential curves of the upper state is observed as pressure broadening of the atomic line.

In the present problem—mercury with an added rare gas—the interaction of the two atoms is simpler because they are different from one another. In

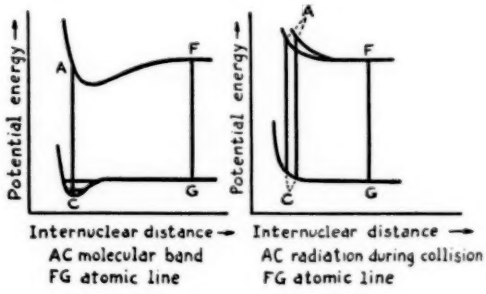

Fig. 3 Fig. 4

Fig. 3. Radiation during collision. AC, radiation during collision; FG, atomic line.

Fig. 4. Potential curves for the mercury and rare gas collision. AC, BD, diffuse maxima on the short wave-length side; AQ, BP, limits of the spectrum; KL, MN band structure on the long wave-length side; FG, atomic line.

the normal state of the two atoms again only one potential curve exists (representing an $^1\Sigma$ term), because both the mercury and the rare gas atom have spherical symmetry in their normal ^1S states. In the upper state (excited atom+normal rare gas atom) the method of counting the molecular electronic levels, given by Hund,[11] depends upon the coupling. In Weizel's problem (He′+He) the multiplet structure (due to the interaction between the orbital momentum and the spin) is narrow as compared with the extension of the continuous spectrum (due to the interaction between the colliding atoms); this interaction between the atoms is described as a quantized orientation of the *orbital momentum* with regard to the internuclear axis. In the present problem, however, it must be taken into account that the splitting up of the multiplet structure in the mercury or thallium atom is much larger than the extension of the diffuse molecular spectrum under consideration. Hence the interaction between the orbital momentum and the spin is large

[11] F. Hund, Zeits. f. Physik **36**, 659 (1926).

as compared with the interaction between the two colliding atoms and a quantized orientation of the *total angular momentum* of the Hg or Tl atom is to be expected with regard to the internuclear axis, just as in the case of a Stark effect in a weak electric field. The number of separate energy levels of the Hg′ or Tl′ colliding with a normal rare gas atom is given by the number of terms in the series $j, j-1 \cdots, 0$ or $\frac{1}{2}$. (Negative values of j describe another orientation but give the same energy levels as the positive values of j).[12] This argument leads to the conclusion that in the upper state Hg′(2^3P_1)+rare gas (^1S) two molecular levels exist, given by $j=1$ and $j=0$, represented in the diagram by two separate potential curves. From both excited levels the transitions to the normal level are allowed. (As in general the collision will not be head on, a rotation will be superimposed with a constant angular momentum.)

Along one or the other of these curves the rare gas atoms and the mercury atoms will approach each other to a certain range of potential energy A or B determined by the thermal energy.

The effective cross section of the Hg atoms in the normal state will be supposed to be smaller than in the excited state. Hence the lowest potential curve of Fig. 4 representing the normal state will just barely begin to rise for the values of nuclear distance that belong to A and B. Therefore both transitions AC and BD, that are vertical in the diagram—following the Franck-Condon rule—represent radiations of larger energy or shorter wave-length than the atomic line FG, BD shifted by a larger amount than AC. This reasoning leads to two separate diffuse maxima on the short wave-length side of the atomic line, both of them shifted by an amount smaller than the mean kinetic energy of the collisions. This just describes the observed spectra. The two separate maxima are interpreted not by the vibration of a molecule but by two separate electronic levels of a pair of atoms to be described by two separate directions of mutual approach, both of them belonging to the same electronic levels of the atoms. Thus the separate maxima give evidence of space quantization in the collision between atoms. This interpretation is strongly supported by the fact that the two separate maxima persist at high temperature (cf. section IV).

The diffuse bands show a rather sharp limit on the short wave-length side and in some cases a much longer diffuse extension to long wave-lengths. This may be understood, too, on the basis of the potential curve diagram (Fig. 4). The dotted lines represent transitions in the case of a small change of internuclear distance during radiation. The long distance BP, representing radiation of short wave-length, cannot be much longer than BD, representing the most probable transition from B. AQ, however, representing the long wave-length extension, can be considerably shorter than AC due to the steep branch of the lower potential curve.[13]

[12] Cf. W. Pauli, Handbuch d. Physik XXIII, p. 246.

[13] The extension to long waves plays the major part in the absorption experiments. at high pressures of Füchtbauer, Joos and Dinckelacker (Ann. d. Physik **71**, 204, (1923)) and Bonhoeffer's and Reichardt's experiments with Hg dissolved in liquids (Zeits. f. Physik **67**, 780 (1931)).

The appearance of the same diffuse structure in the *absorption* spectrum will be explained by the inverse process, the absorption being produced by a mercury atom at the moment of a collision with a rare gas atom.

That these *separate* maxima were not observed by Hopfield in the spectrum of helium, was probably due to their position in the extreme ultraviolet, where the resolution of the same differences in frequency is much more difficult.

The diffuse maxima, reported by Krefft and Rompe in thallium vapor with an added rare gas near the green thallium line $2^2S_{1/2}-2^2P_{3/2}$, are explained by separate molecular levels ($j=\pm 3/2$ and $\pm 1/2$) belonging to the *lower* atomic level instead of the upper one. All other observations, too, can be interpreted on the same basis. One cannot expect, however, to observe separate maxima in every case, for which they are derived from the theory, because they may overlap too much. This is evidently the case in Hg+He, where the existence of two maxima can only be deduced by analogy with the other rare gases.

IV. Quantized Molecules Formed of Excited Mercury and Rare Gas Atoms

This explanation does not take into account the finer diffuse structure on the long wave-length side of the atomic line (Fig. 1). Its previous interpretation as oscillation quanta of a molecule will be maintained.[14] The simple structure and its position on the long wave-length side is most easily explained—contrary to the explanation published previously—by the formation of molecules of Hg atoms in their *excited* state with rare gas atoms. One expects the excited Hg atom to be able to induce an electric moment in the rare gas atom and thus by polarization form a loosely bound molecule. The normal Hg will have this power to a smaller extent because it has a high symmetry similar to that of a rare gas atom. Therefore the upper potential curves of Fig. 4 ought to show shallow minima for large values of the internuclear distance, for which the lower curve is approximately flat. This type of excited molecule would emit a diffuse band on the long wave-length side of the atomic line, in which the vibrational quanta still might be distinguished.

Hence this band system is explained by a group of vibrational levels in the *upper* electronic state combining with a lower state in which vibrational levels can not be resolved. It has been described as an apparently simple set of bands with a trace of convergence to the long wave-length side; it has been mentioned, however, that possibly the observed set of bands is due to the superposition of several sets. According to the present interpretation, two simple sets of bands (KL and MN, Fig. 4) belonging to the two upper potential curves should be superimposed. The smaller slope of the potential curve of the normal state might modify the position of the observed bands. Therefore no conclusion can be drawn concerning the convergence from the structure as shown in Fig. 1.

[14] Cf. F. London, Zeits. f. Physik **11**, 248 (1931).

This interpretation of the diffuse structure on the long wave-length side as vibrational quanta of a band system is supported by observations. This structure does not appear for the light atoms He and Ne, which—due to their smaller polarizability—might not form molecules. (It remains uncertain if this type of structure exists for Hg+X. The pictures with Xenon were taken with a quartz bulb of a few millimeter diameter, from which the scattered light was so strong that the structure might have been concealed.) At high temperature the fine structure on the long wave-length side disappears, while the broad maxima on the short wave-length side remain (observed in Hg+A). This is understood as the complete dissociation by heat; in these circumstances there will remain only the broad maxima on the short wave-length side that are explained not by vibrations of molecules but by separate electronic levels of a pair of colliding atoms.

V. Conclusion

We come to the conclusion that three different processes are to be distinguished in the collision of Hg′ with a rare gas atom. In most cases the atoms will separate without any effect, leaving the Hg′ such that it radiates its sharp atomic line (*FG* in Fig. 4). This may be the most probable process. In some other cases the radiation will take place in the moment of the collision giving rise to the broad maxima on the short wave-length side of the atomic line (*AC* and *BD* in Fig. 4). Finally in a few cases, if the temperature is not too high, some energy may be taken away in a triple collision;[15] real excited molecules with vibrational quanta will be produced emitting a band spectrum (*KL* and *MN* in Fig. 4).

The diffuse maxima in the fluorescence radiation of mercury vapor with a rare gas added are surprisingly intense. The corresponding spectra in the neighborhood of some other lines, for instance the *D* lines of Na, are much fainter or missing entirely. This leads to the question, whether in some atomic terms, but not in all of them, the collision process might produce an instantaneous increase in the probability of emission. This problem will be further investigated.

According to the present interpretation, the main conclusion is that the two separate maxima of intensity, observed in the spectra of certain metal vapors mixed with rare gases, furnish evidence that the mutual approach of these atoms can be described by space quantization.

[15] M. Born and J. Franck, Ann. d. Physik **76,** 225 (1925).

The Equation of State of a Non-ideal Einstein-Bose or Fermi-Dirac Gas

By G. E. UHLENBECK AND L. GROPPER
University of Michigan

(Received May 25, 1932)

With regard to the question if from isotherm measurements one can obtain an experimental test for the existence of Bose statistics in real gases, as is required by theory, we prove the following general theorem. The "Zustandsumme" of a non-ideal Bose or Fermi gas is given by the classical integral provided one replaces the Boltzmann $\exp(-\phi_{ij}/kT)$ factor by:

$$e^{-\phi_{ij}/kT}(1 \pm \exp\left[-4\pi^2 mkTr_{ij}^2/h^2\right] \tag{1}$$

for each pair of molecules (ij). For the second virial coefficient B, this has, i.e., in a Bose gas, as a consequence that:

$$B = B_{\text{non-ideal class}} + B_{\text{ideal Bose}} + B' \tag{2}$$

where:

$$B' = 2\pi N \int_0^\infty drr^2(1 - e^{-\phi(r)/kT}) \exp\left[-4\pi^2 mkTr^2/h^2\right]. \tag{3}$$

Only at very low temperatures do the last two terms in (2) become appreciable. They are then of the same order of magnitude, but have opposite signs. Due to this fact, due to the lack of precise knowledge of the molecular forces, and due to the absence of accurate measurements of B at very low temperatures, one can as yet not decide from isotherm measurements alone whether or not real gases obey the Bose statistics.

I. INTRODUCTION

1.

AS IS well known, it has been proved from the fact that electrons and protons obey the Fermi-Dirac (F.D.) statistics, that a gas consisting of particles of even charge obeys the Einstein-Bose (E.B.) statistics.[1] A neutral gas, for example helium, is a special case of this. The question therefore becomes very important, if one could for example verify this theoretical prediction from equation of state measurements, in particular from the experimentally determined values of the second virial coefficient B.

The theoretical expression B_{cl}, for a *non-ideal* gas, obeying *Boltzmann* statistics is, as is well known:[2]

$$B_{\text{cl}} = 2\pi N \int_0^\infty drr^2(1 - e^{-\phi(r)/kT}) \tag{1}$$

N is the number of molecules per mol. It is here assumed that the intermolecular forces are radial and have the potential $\phi(r)$.

The expression $B_{E.B.}$ or $B_{F.D.}$ for an *ideal* gas obeying Einstein-Bose or Fermi-Dirac statistics, is:

[1] E. Wigner, Sitzungsber. der Ungarischen Acad. 1928, p. 1; P. Ehrenfest and J. R. Oppenheimer, Phys. Rev. **37**, 333 (1931).

[2] See e.g., R. H. Fowler, Statistical Mechanics, Ch. VIII.

$$\left.\begin{matrix} B_{E.B.} \\ B_{F.D.} \end{matrix}\right\} = \mp \frac{1}{2^4}\left(\frac{h^2}{\pi m k T}\right)^{3/2} N \tag{2}$$

In the recent work on the quantum theory of the equation of state of real gases (London, Slater, Kirkwood, Keyes, Margenau) one has made only use of (1), introducing for $\phi(r)$ the theoretical expression obtained from a generalized Heitler-London calculation.[3] The reason that (2) was neglected is, that only at very low temperatures (say below 20°) it becomes appreciable, and *because it was always assumed that the B for a non-ideal Bose gas was the sum of* (1) *and* (2).

We wish to show in this paper that this additivity is not strictly true, but that in addition there occurs a term due to the interaction between the *Boseness* and the non-ideality of the gas. We get for a non-ideal Bose gas:

$$B = B_{cl} + B_{E.B.} + B'_{E.B.} \tag{3}$$

where:

$$B'_{E.B.} = 2\pi N \int_0^\infty dr r^2 (1 - e^{-\phi(r)/kT}) \exp\left[-4\pi^2 m k T r^2/h^2\right] \tag{4}$$

2. The proof of (3) and (4) rests upon a kind of generalization of the $\exp(-V/kT)$ theorem of Boltzmann for an E.B. or F.D. gas. Recently Slater[4] has pointed out the analogue of this fundamental theorem in the quantum theory. He remarked that quantum mechanically, due to the orthogonality and normalization of the wave functions ψ_n belonging to the energy values E_n, the *Zustandssumme* S_Q may be written:

$$S_Q = \sum_n G_n e^{-E_n/kT} = \int \cdots \int dx_1 \cdots dx_N \sum_n e^{-E_n/kT} \psi_n^2. \tag{5}$$

In *Boltzmann-statistics* for a given E_n *all* the G_n linearly independent wave functions ψ_n belonging to it, are allowed, and must therefore occur in the sum,[5] so that the G_n does not explicitly appear any more. In *Bose statistics* the $G_n = 1$, because only those E_n are allowed, to which belong the symmetric wave functions, and only these have therefore to occur in the sum. Similarly in the *F.D. statistics* we have the E_n to which the antisymmetric functions belong. Hereafter we will distinguish between these three cases by writing S_{QB}, $S_{QE.B.}$, $S_{QF.D.}$.

Slater compares (5) now with the classical "Zustandsintegral" after integration over the momenta:

[3] F. London, Zeits. f. Physik **63**, 245 (1930); Zeits. f. phys. Chem. B. **11**, 222 (1930). J. C. Slater, Phys. Rev. **32**, 349 (1928); J. C. Slater and J. G. Kirkwood, Phys. Rev. **37**, 682 (1931). H. Margenau, Phys. Rev. **36**, 1782 (1930); **37**, 1014, 1425 (1931); **38**, 747, 1785 (1931); Proc. Nat. Acad. **18**, 56, 230 (1932). J. G. Kirkwood and F. G. Keyes, Phys. Rev. **37**, 832; **38**, 576 (1931). J. G. Kirkwood, Phys. Zeits. **33**, 39 (1932). One has made small corrections in (1), though, to allow for the existence of discrete states (formation of polarization molecules).

[4] J. C. Slater, Phys. Rev. **38**, 237 (1931).

[5] Comp. P. Ehrenfest and G. E. Uhlenbeck, Zeits. f. Physik **41**, 24 (1927).

$$S_{\text{Cl}} = \left(\frac{2\pi mkT}{h^2}\right)^{3N/2} \int \cdots \int e^{-V/kT} dx_1 \cdots dz_N \tag{6}$$

where $V(x_1 \cdots z_N)$ is the total potential energy due to the intramolecular forces and eventually also due to external forces.

One might expect from the correspondence principle that for sufficient high temperatures where the influence of the discreteness of the translational energy levels can be neglected with respect to the influence of the intermolecular forces (which is already the case, excluding the electron gas, say for $T > 1°$ absolute)[6]

$$\sum_n {}_{B.} e^{-E_n/kT} \psi_n^2 \underset{(T>1°)}{\rightarrow} \left(\frac{2\pi mkT}{h^2}\right)^{3N/2} e^{-V/kT}. \tag{7}$$

Slater makes this also plausible from the meaning of ψ_n^2 as a probability density. In Chapter III we will give an analytical proof. This gives the justification even in the quantum theory for the use of (6) in the case of a Boltzmann gas.

For the case of a E.B. or F.D. gas we will prove a theorem analogous to (7):

$$\left.\begin{matrix} \sum_n {}_{E.B.} e^{-E_n/kT} \psi_n^2 \\ \sum_n {}_{F.D.} e^{-E_n/kT} \psi_n^2 \end{matrix}\right\} \underset{(T>1°)}{\rightarrow} \frac{1}{N!}\left(\frac{2\pi mkT}{h^2}\right)^{3N/2} e^{-V/kT} \times \prod_{ij}(1 \pm \exp\left[-4\pi^2 mkT r_{ij}^2/h^2\right] \tag{8}$$

where the product has to be taken over all pairs of molecules, and r_{ij} is the distance between the (ij) pair. This again holds as indicated for temperatures so high that the influence of the discreteness of the translational energy levels is negligible, but not yet so high that the deviations due to the E.B. or F.D. statistics from the classical gas laws can be neglected. These are given by (8) in *first* approximation; in addition therefore the volume must not be too small. One sees from (8) that there is an apparent attractive force in the E.B. gas between each pair of molecules with the potential:

$$\phi' = -kT \log \left\{1 + e^{-4\pi^2 mkT/h^2 r^2}\right\} \tag{9}$$

Of course for temperatures so high that the deviations due to E.B. or F.D. statistics can be neglected compared to the non-ideality of the gas (T, say for He, $> 20°$ absolute) (8) goes over into (7).

3. Because the total potential energy V (in absence of external fields) can be written in the form:

$$V = \sum_{ij} \phi(r_{ij})$$

(8) can by using (9) be put in the form,

$$\sum_{E.B.} = \frac{1}{N!}\left(\frac{2\pi mkT}{h^2}\right)^{3N/2} \prod_{ij} e^{-(\phi_{ij} + (\phi_{ij}')/kT}. \tag{8a}$$

[6] We will distinguish the integrands of S_{QB}, $S_{QE\cdot B}$ by subscripts B, $E.B.$ etc. and sometimes write for the sums simply Σ_B, $\Sigma_{E\cdot B}$. etc.

Exactly as (7) is carried through by way of (6) to give (1), so (8a) is carried through and will give again (1) when we replace there ϕ by $\phi+\phi'$. Introducing then the value of ϕ' we immediately obtain (3) and (4).

To obtain an idea of the contribution of $B_{E.B.}'$ to the second virial coefficient, we have made in the case of helium a graphical integration, introducing for $\phi(r)$ the theoretical expression used by Kirkwood and Keyes, and the slightly different one proposed by Margenau, as well as the half empirical one calculated by Lennard-Jones.[7] We got:

$a.$ $\qquad T = 20°: B_{E.B.} = -0.8 \quad B'_{E.B.} \leqq +1.2$

$b.$ $\qquad T = 5°: \quad B_{E.B.} = -6.4 \quad B'_{E.B.} \leqq +5.4.$

We see therefore that $B_{E.B.}$ and $B_{E.B.}'$ are of opposite sign and of the same order of magnitude at these low temperatures. At higher temperatures of course both go to zero rapidly. It is clear therefore that because of this fact, it will be still more difficult to decide from the experimentally determined B values at low temperatures, between the classical and the E.B. statistics. In fact, with the uncertainties still existing at present, both in the theory of the intermolecular potential $\phi(r)$ and in the experimental material, in spite of the rather great accuracy and extent of the latter,[8] it seems to us rather hopeless to try to make this distinction with this method.

Quite analogous results are obtained for the F.D. statistics, the signs being merely reversed,

$$B_{F.D.} = -B_{E.B} \qquad B'_{F.D.} = -B'_{E.B.}$$

so that also the existence of the F.D. statistics may not be ruled out.

II. A New Treatment of the Ideal Gas

4. In order to prove (7) and (8), it will be convenient to show it first in the case of the ideal gas, because as we will see in Chapter III, the proof for the general case can be made to rest upon the evaluation of the sum for the ideal gas. Consider first as the simplest case two atoms in a one-dimensional box of length L.

a. With Boltzmann statistics we have to take *all* linearly independent wave functions belonging to each energy value, and we get therefore for the sum:

$$\begin{aligned}\sum_B &= \frac{4}{L^2}\sum_{n_1 \geqq n_2}\sum e^{-\alpha\pi^2/L^2(n_1^2+n_2^2)}\left(\tfrac{1}{2}\right)^{\delta_{n_1 n_2}}\left\{\sin^2\frac{n_1\pi x_1}{L}\sin^2\frac{n_2\pi x_2}{L}\right.\\ &\qquad \left.+\sin^2\frac{n_1\pi x_2}{L}\sin^2\frac{n_2\pi x_1}{L}\right\}\\ &= \frac{4}{L^2}\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} e^{-\alpha\pi^2/L^2(n_1^2+n_2^2)}\sin^2\frac{n_1\pi x_1}{L}\sin^2\frac{n_2\pi x_2}{L}\end{aligned} \tag{10}$$

[7] Comp. R. H. Fowler, Statistical Mechanics, Ch. X.

[8] Comp. G. P. Nijhoff, Dissertation Leiden, 1928.

$$= \left[\frac{2}{L} \sum_{n_1} e^{-\alpha\pi^2/L^2 n_1^2} \sin^2 \frac{n_1 \pi x_1}{L}\right] \cdot \left[\frac{2}{L} \sum_{n_2} e^{-\alpha\pi^2/L^2 n_2^2} \sin^2 \frac{n_2 \pi x_2}{L}\right]$$

where:

$$\alpha = h^2/8\pi^2 mkT$$

In the neighborhood of zero values of x_1 and x_2 we may replace the sums by integrals and obtain:

$$\sum_B \cong \frac{2\pi mkT}{h^2}(1 - e^{-x_1^2/\alpha})(1 - e^{-x_2^2/\alpha}). \tag{11}$$

From the symmetry of the problem it is clear that for x_1 and x_2 in the neighborhood of L we have an analogous expression replacing x_1 and x_2 in (11) by $L - x_1$ and $L - x_2$. We see therefore that Σ_B is zero at the boundary of the square $OABC$ and rises rapidly from all sides to the value $2\pi mkT/h^2$. In the center the wave functions in (10) will average out, because of their different

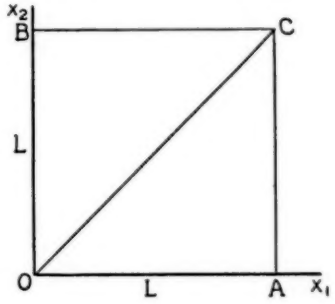

Fig. 1.

phases. Replacing them by the average value 1/4 and the rest of the sum by the integral, (10) again becomes $2\pi mkT/h^2$, so that Σ_B remains constant throughout in accordance with (7). As we will see in §5, the exponentials in (11) account for the correction due to the discreteness of the energy levels.

b. With E.B. or F.D. statistics we get:

$$\left.\begin{matrix}\sum_{E.B.} \\ \sum_{F.D.}\end{matrix}\right\} = \frac{2}{L^2} \sum_{n_1 \geqq n_2} \sum e^{-\alpha\pi^2/L^2(n_1^2+n_2^2)} \left(\tfrac{1}{2}\right)^{\delta_{n_1 n_2}} \left\{\sin \frac{n_1 \pi x_1}{L} \sin \frac{n_2 \pi x_2}{L} \pm \sin \frac{n_1 \pi x_2}{L} \sin \frac{n_2 \pi x_1}{L}\right\}^2. \tag{12}$$

We see that OC is for each term of the sum a symmetry line; in the F.D. case it is a zero line, in the E.B. case a maximum line. This will therefore also be the case for the total sum. We get working the bracket out and going over from the double sum to products of single sums as in *a*:

$$\left.\begin{matrix}\sum_{E.B.} \\ \sum_{F.D.}\end{matrix}\right\} = \tfrac{1}{2} \sum_B \pm \frac{1}{2L^2}\left[\sum_{n=0}^{\infty} e^{-\alpha\pi^2/L^2 n^2} \left\{\cos \frac{n\pi}{L}(x_1 - x_2) - \cos \frac{n\pi}{L}(x_1 + x_2)\right\}\right]^2. \tag{13}$$

All the terms in the first sum of the bracket are in phase along the line $x_1 = x_2$, whereas the second sum has a value only in the corners O and C; in the rest of the square the second cos averages out to zero. Therefore we can neglect the second sum with respect to the first sum. Neglecting also the exponentials in Σ_B (this can really only be justified for N particles, see §5) and replacing the sum by an integral, we get:

$$\left.\begin{matrix}\sum_{E.B.}\\ \sum_{F.D.}\end{matrix}\right\} = \frac{1}{2}\cdot\frac{2\pi mkT}{h^2}\left\{1 \pm e^{-\frac{1}{2\alpha}(x_1-x_2)^2}\right\}. \tag{14}$$

This is in accordance with (8). We have therefore practically everywhere half the classical value; only along OC it rises in the E.B. statistics sharply to the classical value, and in the F.D. statistics it decreases equally sharp to zero.

5.

Going over now to N particles in a one-dimensional box of length L, we have:

a. for Boltzmann statistics:

$$\sum_B \leqq \left(\frac{2\pi mkT}{h^2}\right)^{N/2} \prod_{i=1}^{N}(1 - e^{-x_i^2/\alpha}). \tag{15}$$

This is clearly (11) extended to N particles. From this follows the free energy ψ, according to:

$$e^{-\psi/kT} = \int_0^L \cdots \int dx_1 \cdots dx_N \sum_B. \tag{16}$$

Of course one must remember that (15) holds only for every x_i between 0 and $\frac{1}{2}L$, and that for the other half we have an analogous expression replacing x_i by $L - x_i$. Because of this symmetry we can write Eq. (16):

$$e^{-\psi/kT} = \left(\frac{2\pi mkT}{h^2}\right)^{N/2}\cdot 2^N \int_0^{L/2} \cdots \int dx_1 \cdots dx_N \prod_i (1 - e^{-x_i^2/\alpha}).$$

Of all the terms in the product we need only those which after integration give terms proportional to L^N and L^{N-1}; because we are only interested in the second virial coefficient. We then obtain:

$$e^{-\psi/kT} = \left(\frac{2\pi mkT}{h^2}\right)^{N/2} V^N\left[1 - \frac{N}{V}\left(\frac{h^2}{8\pi mkT}\right)^{1/2}\right] \tag{17}$$

writing V for L now. This gives for the pressure:

$$p = -\frac{\partial\psi}{\partial V} = \frac{NkT}{V}\left[1 + \frac{1}{V}\cdot\frac{1}{2^{3/2}}\left(\frac{h^2}{\pi mkT}\right)^{1/2}\right]. \tag{18}$$

This agrees with the result obtained by the direct calculation of the Zustandssumme.[9] The second term gives the correction due to the discreteness of the

[9] For a gas of N particles in a n-dimensional vessel this gives:

$$p = \frac{NkT}{V}\left[1 + \frac{1}{2^{3/2}}\cdot\frac{1}{V^{1/n}}\cdot\left(\frac{h^2}{\pi mkT}\right)^{1/2}\right]$$

energy levels. This correction is *independent of* N, in contrast to the correction due to the E.B. or F.D. statistics which are *proportional to* N. For this reason the above correction is completely negligible. *Because it arises from the exponentials in* (15) *we will from now on neglect these.*

b. For E.B. or F.D. statistics we get for the sum:

$$\left.\begin{matrix}\sum_{E.B.}\\ \sum_{F.D.}\end{matrix}\right\} = \frac{2^N}{L^N}\frac{1}{N!}\sum_{n_1\geqq n_2\geqq\cdots n_N}\sum\cdots\sum e^{-\alpha\pi^2/L^2(n_1^2+\cdots n_N^2)}g_{n_1\cdots n_N}$$
$$\left[\sum_P{}^{+1}(-1)^P \sin\frac{n_1\pi x_1}{L}\cdots\sin\frac{n_N\pi x_N}{L}\right]^2 \tag{19}$$

where the sum goes over all permutations of the x_i keeping the quantum numbers fixed. The factor $g_{n_1\ldots n_N}$ is one, when all the n_i are different; is $\frac{1}{2!}$ when two are the same and the other different; $\frac{1}{3!}$ when three are the same, etc. When we work the square out we get first of all the square terms, which are just the Boltzmann terms evaluated in *a*. From the rest of the terms we must select only those which after integration over x_i will give terms proportional to L^{N-1}. One can easily convince oneself that these terms are of the type:

$$\sin\frac{n_1\pi x_i}{L}\sin\frac{n_1\pi x_j}{L}\sin\frac{n_2\pi x_i}{L}\sin\frac{n_2\pi x_j}{L}\cdot\begin{cases}\sin^2\dfrac{n_3\pi x_1}{L}\sin^2\dfrac{n_4\pi x_2}{L}\cdots\\ \sin^2\dfrac{n_3\pi x_2}{L}\sin^2\dfrac{n_4\pi x_1}{L}\cdots\\ \cdots\cdots\cdots\end{cases} \tag{20}$$

As in §4 we write for the first four sin:

$$\tfrac{1}{4}\left[\cos\frac{n_1\pi}{L}(x_i-x_j)-\cos\frac{n_1\pi}{L}(x_i+x_j)\right]\times$$
$$\left[\cos\frac{n_2\pi}{L}(x_i-x_j)-\cos\frac{n_2\pi}{L}(x_i+x_j)\right].$$

For the same reason as in §4 we may neglect again the cos of the sum x_i+x_j. Going over from the multiple sum to products of single sums, and replacing the sums by integrals, we get:

$$\left.\begin{matrix}\sum_{E.B.}\\ \sum_{F.D.}\end{matrix}\right\} \leqq \frac{1}{N!}\left(\frac{2\pi mkT}{h^2}\right)^{N/2}\left[1\pm\frac{1}{L^2}\cdot\frac{h^2}{2\pi mkT}\times\right.$$
$$\left.\sum_{ij}\left\{\int_0^\infty e^{-\alpha\pi^2/L^2n^2}\cos\frac{n\pi}{L}(x_i-x_j)dn\right\}^2\right]$$
$$\leqq \frac{1}{N!}\left(\frac{2\pi mkT}{h^2}\right)^{N/2}\left[1\pm\sum_{ij}e^{-(x_i-x_j)^2/2\alpha}\right] \tag{21}$$
$$\leqq \frac{1}{N!}\left(\frac{2\pi mkT}{h^2}\right)^{N/2}\prod_{ij}\left\{1\pm e^{-(x_i-x_j)^2/2\alpha}\right\}.$$

The sum and product are to be taken over all pairs. The last two lines differ only in terms which after integration over x_i give terms proportional to powers of L lower than $N-1$, and therefore they have been neglected. Formula (21) agrees with (8), when we put there $V=0$.

To check (21) further we may calculate the free energy. One obtains:

$$e^{-\psi/kT} = \int_0^L \cdots \int dx_1 \cdots dx_N \begin{Bmatrix} \sum_{E.B.} \\ \sum_{F.D.} \end{Bmatrix}$$
$$= \frac{1}{N!}\left(\frac{2\pi mkT}{h^2}\right)^{N/2} V^N \left[1 \pm \frac{1}{V}\,\frac{N(N-1)}{2}\left(\frac{h^2}{4\pi mkT}\right)^{1/2}\right]$$

where again we write V for L. Neglecting the 1 with respect to N, we get for the pressure:

$$p = -\frac{\partial\psi}{\partial V} = \frac{NkT}{V}\left[1 \mp \frac{N}{V}\cdot\frac{1}{2^2}\left(\frac{h^2}{\pi mkT}\right)^{1/2}\right] \tag{22}$$

which agrees with well known results.[10] Comparing (15) and (21), and their consequences (18) and (22) resp. one sees that the reason why the E.B. or F.D. correction has the factor N, in contrast to the discreteness correction, is simply that the number of pairs, which one can form with N objects (x_i), is $\frac{1}{2}N(N-1)$.

These considerations can immediately be extended to a gas in a three-dimensional rectangular vessel. One obtains then instead of (21):

$$\left.\begin{matrix} \sum_{E.B.} \\ \sum_{F.D.} \end{matrix}\right\} = \frac{1}{N!}\left(\frac{2\pi mkT}{h^2}\right)^{3N/2} \prod_{ij}(1 \pm e^{-(1/2\alpha)[(x_i-x_j)^2+(y_i-y_j)^2+(z_i-z_j)^2]} \tag{21a}$$

so that one sees that the apparent potential due to the Bose-ness is radial (see Eq. (9)).

III. The Non-Ideal E.B. or F.D. Gas

§6.

The wave equation for the whole gas we write as:

$$(H - E)\psi = 0$$

where:

$$\left.\begin{aligned} H &= \sum_i \frac{1}{2m}(p_{xi}^2 + p_{yi}^2 + p_{zi}^2) + V(x_1 \cdots z_n) \\ V &= \sum_{ij}\phi_{ij} + \sum_i \pi_i \end{aligned}\right\} \tag{23}$$

where ϕ_{ij} is the interaction potential between the particles and π_i an eventual external potential. We shall abbreviate always all the coordinates by q and all the quantum numbers by n. We wish to show in this paragraph that:

[10] For an E.B. or F.D. gas in an n-dimensional vessel one finds according to the usual method:

$$p = \frac{NkT}{V}\left[1 \mp \frac{1}{2^{n+1}}\left(\frac{h^2}{\pi mkT}\right)^{n/2}\frac{N}{V}\right].$$

$$\sum_n e^{-E_n/kT}\psi_n^2 = \sum_n \phi_n e^{-H/kT}\phi_n. \tag{24}$$

In the right-hand side H is to be considered as an operator working on ϕ_n, which are the normalized eigenfunctions for the *ideal* gas.[11] Formula (24) is true *regardless of the statistics*, that is to say, it holds equally well when we take on both sides only the symmetric wave functions (E.B. statistics), or only the antisymmetric wave functions (F.D. statistics), or all the wave functions (Boltzmann statistics).

For the proof one remarks first that for any function (power series) $f(x)$, we have:

$$f(E_n)\psi_n = f(H)\psi_n$$

in which H again is the operator, in particular:

$$e^{-E_n/kT}\psi_n = e^{-H/kT}\psi_n. \tag{25}$$

We develop now ψ_n according to the normalized eigenfunctions of the ideal gas. In the case of a cubicle vessel this is simply a development in a multiple Fourier-series

$$\psi_n(q) = \sum_m c_{nm}\phi_m(q) \tag{26}$$

where one must then remember that $\phi_m(q)$ is an abbreviation for:[12]

$$\phi_{m_1\cdots m_{3N}}(x_1 \cdots z_N) = \left(\frac{2}{L}\right)^{3N/2} \sin\frac{m_1\pi x_1}{L}\sin\frac{m_2\pi y_1}{L}\cdots\sin\frac{m_{3N}\pi z_N}{L}.$$

Substituting (25) and (26) in the left-hand side of (24) we get:

$$\sum_n \psi_n e^{-E_n/kT}\psi_n = \sum_n \sum_i \sum_j c_{ni}c_{nj}\phi_i e^{-H/kT}\phi_j = \sum_i \phi_i e^{-H/kT}\phi_i$$

since: $\Sigma_n c_{ni}c_{nj} = \delta_{ij}$.

§7.

Introducing for simplification:

$$V' = 8\pi^2 mV/h^2 \quad E' = 8\pi^2 mE/h^2$$

the wave equation becomes:

$$d^2\psi/dq^2 + (E' - V')\psi = 0$$

and for the sum, using (24) we get:

$$\sum_n e^{-E_n/kT}\psi_n^2 = \sum_n e^{-\alpha E_n'}\psi_n^2 = \sum_n \phi_n e^{\alpha(d^2/dq^2 - V')}\phi_n$$

[11] Comp. F. Bloch, Zeits. f. Physik **74,** 295 (1932). We came to this result independently and because our point of view is somewhat different, we thought it not superfluous to repeat the proof.

[12] This is in the case of the Boltzmann statistics. For the E.B. or F.D. statistics one takes the symmetric or antisymmetric combinations, and because these form with respect to symmetric resp. antisymmetric functions a *complete* orthogonal set, the argument remains the same.

where α has again the meaning $h^2/8\pi^2 mkT$ and where q abbreviates again the $3N$ coordinates of the gas.

We will prove now that *when α is very small*, we have *in first approximation*:

$$\sum_n \phi_n e^{\alpha(d^2/\alpha q^2 - V')}\phi_n \leqq e^{-\alpha V'} \sum_n e^{-\alpha \epsilon_n'}\phi_n^2 \tag{27}$$

where $h^2\epsilon_n'/8\pi^2 m$ are the energy values of the ideal gas. In particular for a cubicle vessel

$$\epsilon_n' = \pi^2(n_1^2 + n_2^2 + \cdots n_{3N}^2)/L^2.$$

In this way the evaluation of the sum for the *non-ideal* gas reduces to that for the *ideal* gas. Using then the results of Chapter II, we will then have proved our main theorems (7) and (8).

Formula (27) is not obvious, because d^2/dq^2 and V' are not commutable, so that in general:

$$e^{\alpha(d^2/dq^2 - V')} = e^{\alpha(-V' + d^2/dq^2)} \neq e^{-\alpha V'} \cdot e^{\alpha d^2/dq^2}$$
$$\neq e^{\alpha d^2/dq^2} \cdot e^{-\alpha V'}.$$

To prove (27), write:

$$F = e^{\alpha(d^2/dq^2 - V')}\phi_n$$

then F will fulfill the differential equation:[13]

$$\partial F/\partial\alpha = (\partial^2/\partial q^2) - V')F$$

which we rewrite:

$$\partial F/\partial\alpha + (\epsilon_n' + V')F = (\partial^2/\partial q^2 + \epsilon_n')F \tag{28}$$

For $\alpha = 0$, $F = \phi_n$ and the right member is zero. In *first* approximation therefore:

$$F = \phi_n e^{-\alpha(\epsilon_n' + V')} \tag{29}$$

which gives (27). In Note I we will prove that in higher approximations:

$$\sum_n \phi_n e^{\alpha(d^2/dq^2 - V')}\phi_n = e^{-\alpha V'} \sum_n e^{-\alpha\epsilon_n'}\phi_n^2 \cdot [1 + \alpha f_1(q) + \alpha^2 f_2(q) + \cdots] \tag{30}$$

which justifies more strictly that (27) is true for small α; $f_1(q)$, $f_2(q)$ are certain functions of the coordinates, which contain V and its derivatives.

As a special example, which is of interest because the summations can be carried out exactly, we will consider in Note II the case of the harmonic oscillator.

Note I.

To find the next approximation we substitute (29) in the right-hand side of (28). Considered as a differential equation for F as a function of α, we can immediately integrate. Making use of the Schrödinger equation:

[13] Comp. F. Bloch, reference 11.

$$(d^2/dq^2 + \epsilon_n')\phi_n = 0$$

we obtain:

$$F = \left[\phi_n - \frac{\alpha^3}{2}\left\{\phi_n\frac{d^2V'}{dq^2} + 2\frac{d\phi_n}{dq}\cdot\frac{dV'}{dq}\right\} + \frac{\alpha^3}{3}\phi_n\left(\frac{dV'}{dq}\right)^2\right]e^{-\alpha(\epsilon_n' + V')} \tag{31}$$

Substituting in the left-hand side of (27) and calling the sum for the ideal gas:

$$S = \sum_n e^{-\alpha\epsilon_n}\phi_n^2$$

we get:

$$\sum_n \phi_n e^{\alpha(d^2/dq^2 - V')}\phi_n \cong e^{-\alpha V'}\left\{S - \frac{\alpha^2}{2}\left(\frac{d^2V'}{dq^2}S + \frac{dV'}{dq}\cdot\frac{dS}{dq}\right) + \frac{\alpha^3}{3}\left(\frac{dV'}{dq}\right)^2\cdot S\right\}$$

Substituting the value of S from (15) and (21), only the third term of the right-hand member will give a term proportional to α. Neglecting the terms which give α^2 and α^3, we obtain:

a. for Boltzmann statistics:

$$\sum_B = \left(\frac{1}{4\pi\alpha}\right)^{3N/2} e^{-\alpha V'}\prod_i(1 - e^{-x_i^2/\alpha})\cdot\left[1 - \alpha\sum_i \frac{\partial V}{\partial x_i}\,\frac{x_i e^{-x_i^2/\alpha}}{1 - e^{-x_i^2/\alpha}} + \cdots\right] \tag{32}$$

b. for E.B. or F.D. statistics:

$$\left.\begin{matrix}\sum_{E.B}\\ \sum_{F.D}\end{matrix}\right\} = \frac{1}{N'}\left(\frac{1}{4\pi\alpha}\right)^{3N/2} e^{-\alpha V'}\prod_{ij}(1 \pm e^{-(x_i - x_j)^2/2\alpha)})$$
$$\cdot\left[1 \pm \frac{\alpha}{4}\sum_{ij}\frac{(x_i - x_j)e^{(-1/2\alpha)(x_i - x_j)^2}}{1 \pm e^{-(x_i - x_j)^2/2\alpha}}\left(\frac{\partial V'}{\partial x_i} - \frac{\partial V'}{\partial x_j}\right) + \cdots\right]. \tag{33}$$

We must remark, that the terms proportional to α in (32) and (33) contribute to the second virial coefficient, but as one easily verifies they will give corrections to (3) which are proportional to α^3 (for a gas in a three dimensional vessel), whereas (2) is only proportional to $\alpha^{3/2}$.

Further one must remark that the second term in the bracket of (32) and (33) are *not* the $f_1(g)$ of (30), for if one proceeds to the next approximation, one gets another term proportional to α, which contains $\partial^2 V'/\partial x_i^2$. One can convince oneself though that in the second virial coefficient these terms give corrections proportional to powers of α higher than α^3.

One can approach the proof of (27) from a somewhat more general standpoint. When f and g are any two non-commutable quantities, one can show that for small values of α:[14]

$$e^{\alpha(f+g)} = \{1 + (\alpha^2/2)(gf - fg) + \cdots\}e^{\alpha f}\cdot e^{\alpha g}$$

When we identify now g with d^2/dg^2 and f with V', and let $\exp[\alpha(f+g)]$ work on ϕ_n one obtains again (31).[15]

Note II.

For the harmonic oscillator it is possible to carry out the sums exactly. Introducing the abbreviations:

$$y = 2\pi x(m\nu/h)^{1/2} \qquad\qquad \theta = h\nu/kT$$

we can write:

$$(2\pi mkT)^{-1/2}\sum_n e^{-E_n/kT}\psi_n^2 = (2\theta)^{1/2}e^{-\theta/2 - y^2}\sum_n e^{-n\theta}\frac{H_n^2(y)}{2^n\cdot n!} \tag{34}$$

[14] Comp. J. E. Campbell, Theory of continuous Groups, pp. 54–57, for analogous questions.

[15] One uses here the special case $n = 2$ of a general formula for

$$f\frac{d^n}{dg^n} - \frac{d^n}{dg^n}\cdot f.$$

derived by Hylleras (Zeits. f. Physik **74,** 216 (1932)).

where ν is the frequency and H_n the n^{th} Hermite polynomial. Introducing Weber's function $D_n(y)$:[16]

$$H_n(2^{-1/2}y) = 2^{n/2}e^{y^2/4}D_n(y)$$

the right-hand side becomes:

$$(2\theta)^{1/2}e^{-\theta/2}\sum_n e^{-n\theta}D_n^2(2^{1/2}y)/n!.$$

We distinguish again *two* cases;

a. Boltzmann statistics. We have to sum over *all* values of n. From the general result:[17]

$$\sum_n \frac{D_n(\xi)D_n(\eta)}{n'}e^{-nx} = (1-e^{-2x})^{-1/2}e^{((\xi^2+\eta^2))/4}\exp\left\{-\frac{\xi^2+\eta^2-2\xi\eta e^{-x}}{2(1-e^{-2x})}\right\} \tag{35}$$

putting $x=\theta$; $\xi=\eta=2^{1/2}y$, we obtain:

$$(2\pi mkT)^{-1/2}\sum_n e^{-E_n/kT}\psi_n^2 = \left(\frac{2\theta}{1-e^{-2\theta}}\right)^{1/2}\exp\left(-\frac{\theta}{2}-y^2\operatorname{tgh}\frac{\theta}{2}\right). \tag{36}$$

One sees that for θ very small the right-hand side goes over into:

$$e^{-\theta y^2/2} = e^{-2\pi^2 m\nu^2x^2/kT} = e^{-V/kT}.$$

b. E.B. or F.D. statistics. We must suppose then that we have to do with two particles attracting each other with a force proportional to the distance between them, moving in a one-dimensional vessel. If x stands for the distance between the particles, in E.B. statistics we must take only the eigenfunctions *even* in x, and in F.D. statistics we must take the *odd* ones. Separating into center of gravity and relative motion coordinates,[18] we have then to consider for the relative motion simply the sum (34) where in E.B. statistics we have to sum over all even values of n, and in F.D. statistics over the odd values[19] and where one must remember to replace m by $\frac{1}{2}m$. Using now the general result;[20]

$$\sum_{n\text{even}}^{\sum}\left\{\frac{D_n(\xi)D_n(\eta)}{n'}e^{-nx} = (1-e^{-2x})^{-1/2}\exp\left(-\frac{\xi^2+\eta^2}{4}\operatorname{cotgh}x\right)\begin{matrix}\cosh\\ \sinh\end{matrix}\left\{\frac{\xi\eta e^{-x}}{1-e^{-2x}}\right\}\right. . \tag{37}$$

From this follows then immediately:

$$\left.\begin{matrix}(\pi mkT)^{-1/2}\sum_{\text{E.B.}}\\ (\pi mkT)^{-1/2}\sum_{\text{F.D.}}\end{matrix}\right\} = \tfrac{1}{2}\left(\frac{2\theta}{1-e^{-2\theta}}\right)^{1/2}\exp\left(-\frac{\theta}{2}-y^2\operatorname{tgh}\frac{\theta}{2}\right)\cdot\left\{1\pm\exp\left(-\frac{2y^2}{\sinh\theta}\right)\right\} \tag{38}$$

For small θ the right-hand side goes over into:

$$\tfrac{1}{2}e^{-\pi^2m\nu^2x^2/kT}(1\pm e^{-4\pi^2mkTx^2/h^2})$$

in accordance with (8).

[16] Whittaker-Watson, Modern Analysis, p. 347.

[17] See G. E. Uhlenbeck and L. S. Ornstein, Phys. Rev. **36,** 823 (1930). F. Bloch reference 11 also gives the result, using a transformation function of Kennard (Zeits. f. Physik **44,** 326, 1927.)

[18] Strictly this is only possible for an infinite vessel. Because we are not interested here in the discreteness effect of the total translational motion, this is quite allowed.

[19] The center of gravity motion will simply give an extra factor $(4\pi mkT/h^2)^{1/2}$.

[20] The proof is quite analogous as given for (35) by G. E. Uhlenbeck and L. S. Ornstein, reference 17.

On Ferromagnetism and Related Problems of the Theory of Electrons*

By PAUL S. EPSTEIN
California Institute of Technology

(Received May 20, 1932)

In this paper is presented a study of those electrons which hold together a chemically elementary crystal and are responsible for the homoeopolar bonds between neighboring atoms. The starting point of the theory is Slater's form of the secular equations for first order perturbations arising from the interactions of identical atoms composing a system. A rigorous solution of these equations is given in section 4. The investigation is carried through in a unified way for magnetic and non-magnetic materials, with respect to internal energies and to magnetic properties. As anticipated by Heisenberg, the deciding factor is the sign of the Heitler and London interchange integral J_1. Substances with large negative J_1 are non-magnetic, those with large positive J_1, potentially ferromagnetic. For non-magnetic bodies, the theory gives a confirmation of Bloch's conclusions with slight differences of interpretation. On the other hand, the results for magnetic materials are new and entirely different from Bloch's. As to specific heats, it is found that at very low temperatures they have the expression $c=0.208\ sR(T/\theta)^{3/2}$, where R is the gas constant, s the number of valency electrons per atom, and θ has a close relation to the Curie point. With respect to ferromagnetism, the result is that a crystal satisfying Slater's equations is spontaneously magnetized almost to saturation but that the polarity of this magnetization changes its sense at irregular intervals. This fact suggests that ferromagnetic crystals must have a block structure and that they are coherent, in the sense of the validity of Slater s equations, only within the blocks (compare sections 10 and 11). With this hypothesis the theory accounts for the fundamental facts of ferromagnetism.

1. INTRODUCTION

THE discoveries of quantum dynamics, especially Pauli's exclusion principle and Fermi's statistics based on it, gave a new stimulus to the theory of electrons in metals. While Sommerfeld[1] treated in an exhaustive way the subject of free electrons, it was pointed out by Heisenberg[2] that the exclusion principle must exercise a powerful influence on the orientation of the spins of valency electrons and that it may play the role of the *molecular field* postulated by Weiss for ferromagnetic materials. The theory of the more tightly bound electrons was investigated by Bloch in a series of important papers. While his work is fundamental in many respects, it treats the non-magnetic properties[3] and the magnetic[4] separately and by different methods. Moreover, we believe that his results relating to ferromagnetism are open to objection.

* An abridgment of this paper was read before the *National Academy of Sciences* on April 26, 1932 at Washington, D. C.

[1] A. Sommerfeld, Zeits. f. Physik **47**, 1, 43 (1928).

[2] W. Heisenberg, Zeits. f. Physik **49**, 619 (1928).

[3] F. Bloch, Zeits. f. Physik **52**, 555 (1928); **59**, 208 (1930).

[4] F. Bloch, Zeits. f. Physik **57**, 545 (1929); **61**, 206 (1930); **74**, 295 (1932).

[5] This assumption is made for simplicity but is not necessary. We shall see in section 5 that our theory applies, almost without change, to more general cases.

2. Formulation of the Mathematical Problem

We consider with Heisenberg and Bloch a system of N electrons each pertaining, in the unperturbed state, to a different atom of a crystal.[5] Let the orbital motion of the electron be that corresponding to the lowest quantum state. Besides, with every electron there is associated a spin the projection of which on a given axis can be either positive or negative. Let r electrons have negative spins and $N-r$ positive. The projection of the total spin moment is then equal to $mh/2\pi$,where $2m=N-2r$. If all the atoms are successively labeled by the numbers 1, 2, $\cdots N$, let the positions of those containing an electron with a negative spin be $f_1, f_2, \cdots f_r$ $(f_1<f_2<\cdots<f_r)$. We denote the antisymmetric wave function for such a distribution by $\psi(f_1, f_2, \cdots f_r)$. As the magnetic interactions are neglected, all the functions with the same number r, correspond to the same energy and the total wave function pertaining to the magnetic moment m can be written as the linear expression

$$\psi(r) = \sum_{f_1,\cdots f_r} a(f_1, \cdots f_r)\psi(f_1, \cdots f_r). \tag{1}$$

The coefficients $a(f_1, \cdots f_r)$ are determined, in the usual way, by a secular system of equations when we introduce the mutual influence of the electrons and protons as a perturbation. The form of the secular equations was explicitly derived by Slater:[6]

$$\epsilon a(f_1, \cdots f_r) + \sum J_{ff'}[a(f_1', \cdots f_r') - a(f_1, \cdots f_r)] = 0. \tag{2}$$

The summation is extended over all pairs of electrons having opposite spins and $f_1', \cdots f_r'$ and $f_1, \cdots f_r$ differ only in that this pair of spins is interchanged. The coefficient $J_{ff'}$ denotes the London and Heitler interchange integral for the pair and ϵ is the energy measured from an arbitrary zero point. Usually, it is permissible to neglect the mutual influence of the more distant atoms and to take into consideration only the interactions between contiguous electrons. In this case the Eq. (2) is for most crystals reduced to

$$\epsilon a(f_1, \cdots f_r) + J_1 \sum[a(f_1', \cdots f_r') - a(f_1, \cdots f_r)] = 0. \tag{3}$$

The sum includes now all pairs of contiguous electrons having opposite spins. J_1 represents the interchange integral for two adjacent atoms.

The main problem is to find a solution of the Eqs. (2) or (3), satisfying the conditions at the surface of the crystal and other subsidiary conditions which we shall state in the next section. Bloch has given a very elegant method of transforming these equations. However, we shall not use it because the solution can be derived directly from the form given here.

3. Solution in a Special Case

In order to show how the solution is arrived at, it is best to start from an example. We take, therefore, the case of a linear chain of atoms with $r=2$. The form which Eq. (3) takes in this case is

[6] J. C. Slater, Phys. Rev. **34,** 1293 (1929); **35,** 509 (1930).

$$\begin{aligned}\epsilon a(f_1, f_2) + J_1[&a(f_1 - 1, f_2) - a(f_1, f_2) + a(f_1 + 1, f_2) - a(f_1, f_2)\\ &+ a(f_1, f_2 - 1) - a(f_1, f_2) + a(f_1, f_2 + 1) - a(f_1, f_2)] = 0,\end{aligned} \tag{4a}$$

which holds for $f_2 > f_1 + 1$. In the case $f_2 = f_1 + 1$ two of the interchanges are precluded and the equation becomes

$$\epsilon a(f_1, f_2) + J_1[a(f_1 - 1, f_2) - a(f_1, f_2) + a(f_1, f_2 + 1) - a(f_1, f_2)] = 0. \tag{4b}$$

The difference equation (4a) has constant coefficients and its particular solutions can be written in the form $\exp i\alpha(k_1 f_1 + k_2 f_2)$. The corresponding energy level is

$$\epsilon_{k_1, k_2} = 4J_1(\sin^2 \tfrac{1}{2}\alpha k_1 + \sin^2 \tfrac{1}{2}\alpha k_2). \tag{5}$$

This expression is symmetrical in k_1 and k_2 so that the solution $\exp i\alpha(k_2 f_1 + k_1 f_2)$ belongs to the same energy. We can, therefore, build up the slightly more general solution $a(k_1, k_2) = c_1 \exp i\alpha(k_1 f_1 + k_2 f_2) + c_2 \exp i\alpha(k_2 f_1 + k_1 f_2)$ and try to choose the coefficients c_1 and c_2 in such a way as to satisfy the subsidiary condition (4b). It is important to note that while (4a) is supposed to apply only in the case $f_2 > f_1 + 1$, our solution is of such a type that it satisfies the Eq. (4a) for all values of f_1 and f_2, including $f_2 = f_1 + 1$. In this last case, both Eqs. (4a) and (4b) are true and we can replace the subsidiary condition by the difference

$$\begin{aligned}a(f_1 + 1, f_2) + a(f_1, f_2 - 1) - 2a(f_1, f_2) = 0,\\ f_2 = f_1 + 1.\end{aligned} \tag{6}$$

It is easy to see that this condition is satisfied if we put

$$c_1 : c_2 = -\exp i[\alpha(k_1 - k_2) + 2\phi(k_1, k_2)],$$

where

$$\phi(k_1, k_2) = \operatorname{arctg} \frac{\sin \alpha k_1 - \sin \alpha k_2}{2 - \cos \alpha k_1 - \cos \alpha k_2}. \tag{7}$$

The solution itself can then be written in the form

$$\begin{aligned}a(k_1, k_2) = \exp i[&\alpha(k_1 f_1 + k_2 f_2) + \alpha k_1 + \phi(k_1, k_2)]\\ &- \exp i[\alpha(k_2 f_1 + k_1 f_2) + \alpha k_2 + \phi(k_2, k_1)].\end{aligned} \tag{8}$$

If we substitute for the surface conditions the periodicity requirement of Bloch's, all we have to do is to define k_1, k_2 as integers and $\alpha = 2\pi/N$. The Eqs. (8) and (5) will then represent a rigorous solution of the problem. It differs from Bloch's result only inasmuch as $a(k_1, k_2)$ vanishes for $k_1 = k_2$. This means that the case of two equal quantic numbers does not occur and must be excluded.

To obviate possible objections, we prefer, however, to show how a solution satisfying the actual border conditions can be found. These conditions result from the fact that when an electron is at one of the ends of the chain ($f_1 = 1$ or $f_2 = N$) a further decrease (or increase) of the number f labelling its position is no longer possible. The corresponding transitions must be, therefore, excluded from the conditions (4a) and (4b). This means

$$a(f_1 - 1, f_2) - a(f_1, f_2) = O, \qquad \text{when } f_1 = 1, \tag{9a}$$

$$a(f_1, f_2 + 1) - a(f_1, f_2) = O, \qquad \text{when } f_2 = N. \tag{9b}$$

It is easy to see that in addition to $a(k_1,k_2)$ the expressions $a(-k_1, k_2)$, $a(k_1, -k_2)$, $a(-k_1, -k_2)$ are also solutions of the system (4a), (4b) pertaining to the same energy (5), because this energy does not change if we reverse the sign of k_1 or k_2. Combining these four expressions with suitable coefficients we can build up a solution satisfying the condition (9a). In fact, the Eq. (9a) is fulfilled if the expression is arranged as a sum of terms each containing the factor $\cos\alpha k_j(f_1-\frac{1}{2})$, where j is 1 or 2. Such an arrangement is possible for any value of the parameters αk_j. On the other hand, the border condition (9b) can be satisfied by the same expression only for definite and discreet values of these parameters. It is obvious from the symmetry with respect to the two ends of the chain that α must be taken as

$$\alpha = \frac{\pi}{N}. \tag{10}$$

The condition (9b) is equivalent to the requirement that the solution can be arranged as a sum of terms each containing the factor $\cos\alpha k_j\,(f_2-N-\frac{1}{2})$. The arguments of the cos functions can be reduced to this form only if we impose upon the numbers k_j the conditions

$$\left.\begin{aligned}(1 - 1/N)k_1 - (1/\pi)[\phi(k_1, k_2) + \phi(k_1, -k_2)] &= k_1',\\ (1 - 1/N)k_2 - (1/\pi)[\phi(k_2, k_1) + \phi(k_2, -k_1)] &= k_2'.\end{aligned}\right\} \tag{11}$$

Where the numbers k_1', k_2' are integers $(1, 2, \cdots N)$ and represent the quantic numbers of the problem. The rigorous border conditions lead, therefore, to a different result from Bloch's periodicity relation: The numbers k_j are no longer quantic numbers but non-integral auxiliary parameters.

The coefficients of the secular problem acquire in this case the simple form

$$a(f_1, f_2) = \cos\alpha k_1(f_1 - \tfrac{1}{2})\cos\alpha k_2(f_2 - N - \tfrac{1}{2}) \mp \cos\alpha k_2(f_1 - \tfrac{1}{2})\cos\alpha k_1(f_2 - N - \tfrac{1}{2}).$$

4. General Solution

The results of the preceding section apply with slight changes to the general case. For a system with $r=\frac{1}{2}N-m$ negative spins the energy expression is quite generally

$$\epsilon_m = \sum_{j=1}^{\frac{1}{2}N-|m|} \epsilon_j. \tag{12}$$

In the case of a linear chain, we have

$$\epsilon_j = 4J_1 \sin^2 \frac{\alpha k_j}{2N}, \tag{13}$$

where the parameters k_j are determined in relation to the quantic numbers $k_j'(=1, 2, \cdots N)$ by the equations

$$\left(1-\frac{1}{N}\right)k_j-\frac{1}{\pi}\sum_{i\neq j}i[\phi(k_j, k_i)+\phi(k_j, -k_i)]=k_j', \qquad i, j=1, 2, \cdots, \tfrac{1}{2}N-|m|. \tag{14}$$

The expression for $a(f_1, \cdots f_r)$ becomes rather cumbersome. It is built up of terms derived from the product

$$\cos \alpha k_1(f_1-\tfrac{1}{2})\cdot\cos\left[\alpha k_2(f_2-3/2)-\phi(k_2, k_1)-\phi(k_2, -k_1)\right]$$
$$\cdot\cos\left[\alpha k_3(f_3-3/2)-\phi(k_3, k_1)-\phi(k_3, -k_1)-\phi(k_3, k_2)-\phi(k_3, -k_2)\right]\cdots$$

by all possible permutations of $k_1, k_2, \cdots k_r$.

If the number N is large, the difference between k_j and k_j' does not materially affect the energy expressions for smaller values of r, i.e., in the neighborhood of magnetic saturation. As to the opposite case $r=N/2$, $m=0$, the terms of the energy expression in which k_j is small and those in which it is large (near to N) remain practically unchanged. But the middle range of values of k_j, about $N/2$, is considerably influenced in the sense of an increase of this number. In the more important of our applications the states of small magnetic moment will play but a negligible role. Therefore, we shall be satisfied with the approximation which is obtained by identifying k with k', as it is amply sufficient for our purpose.

For a two dimensional quadratic lattice

$$\epsilon_j=4J_1\left(\sin^2\frac{\pi k_j}{2G}+\sin^2\frac{\pi l_j}{2G}\right), \qquad G^2=N. \tag{15}$$

In the case of a three-dimensional cubical crystal

$$\epsilon_j=4J_1\left(\sin^2\frac{\pi k_j}{2G}+\sin^2\frac{\pi l_j}{2G}+\sin^2\frac{\pi n_j}{2G}\right), \qquad G^3=N. \tag{16}$$

The parameters k_j, l_j, n_j are in both cases connected with a set of quantic numbers k_j', l_j', $n_j'=1, 2, \cdots G$ by the relations (14) (with G in place of N) and by two similar sets of equations which are obtained from (14) by substituting l or n instead of k.

In all cases, the Eq. (13) can be divided into a main condition, analogous to (4a), which holds when no two negative spins are in contiguous atoms, and into subsidiary conditions like (4b) which are true in the opposite case. The main condition is satisfied by an exponential function (analogous to $\exp\alpha(k_1f_1+k_2f_2)$ of the preceding section) and by a large number of other exponentials derived from this. The form of the energy terms in the sum (12) is here also completely defined by the main condition and by the exponential function satisfying it. The subsidiary conditions are satisfied by a proper combination of the exponentials and do not affect the form of the energy expressions. Finally, the border conditions serve to define the parameters entering into

the energy terms. Bloch succeeded in solving the main condition, therefore, his energy expressions are formally identical with ours, although the symbols have a somewhat different meaning. To this extent his results for other types of two and three dimensional lattices can also be used. However, if we inquire into the substance instead of the form, there is a profound difference in two other respects which are far more important than the changed definition of the parameters.

(1) In the first place, *no two terms of the energy sum* (12) *may be equal.* In the example of section (3) we see from the formula (8) that $a(f_1, f_2)$ vanishes when $k_1 = k_2$. This means that the corresponding energy level does not exist. The conditions are the same in the general case. Whenever two triples of numbers k_j, l_j, n_j are equal the function $a(f_1, \cdots f_r)$ vanishes identically. This is by no means surprising because the exclusion principle is contained in the formulation of Slater's Eqs. (2) or (3) and makes itself manifest in all the conclusions drawn from it.

(2). In the second place, the number of terms in the sum (12) depends only on the absolute value of the magnetic moment. It is equal to $\frac{1}{2}N - |m|$, and not to $\frac{1}{2}N - m$ as Bloch has it. The mathematical reason for this is that the case of no two negative spins being in contiguous atoms can be realized only as long as the number of these negative spins is smaller than (or equal to) half the number of atoms ($r = \frac{1}{2}N - m \leqq \frac{1}{2}N$, or $m \geqq 0$). If this requirement is not satisfied, we have no main condition but only subsidiary conditions, which completely changes the aspect of the problem. However, we can reduce the new problem to the old one by reversing the direction of our system of coordinates, and by counting as negative those spins which were considered as positive. *To positive and negative magnetic moments of the same absolute magnitude correspond the same energy levels.* This fact is an immediate consequence of the fundamental assumptions of this theory (section 2) which neglects magnetic interactions. Both directions in space are equivalent and the energy cannot depend on sense of the magnetic moment.

These two features of our energy expressions have an important influence on the results and lead to conclusions which are considerably different from Bloch's. Before closing this section, we should like to point out that the theory can be carried through also in the case of Slater's general Eq. (2) which takes into account interactions between non-contiguous atoms. For instance, in the case of a cubical crystal the energy expressions become

$$\epsilon_j = 4 \sum_{g_1 g_2 g_3} J_{g_1 g_2 g_3} [\sin^2 \tfrac{1}{2}\alpha g_1 k_j + \sin^2 \tfrac{1}{2}\alpha g_2 l_j + \sin^2 \tfrac{1}{2}\alpha g_3 n_j]. \qquad (18)$$

where $J_{g_1 g_2 g_3}$ is the interchange integral for two atoms with a distance between them having the projections $g_1 d$, $g_2 d$, $g_3 d$ if d denotes the lattice constant.

5. The Sum of States

To the energy expressions (12) there must be added the magnetic energy of the crystal equal to $-2\mu m H$, where μ denotes the Bohr magneton and H the strength of the magnetic field. The total energy is, therefore,

$$\epsilon = \epsilon_m - 2\mu m H. \tag{19}$$

The part ϵ_m depends only on the absolute value of the moment so that $\epsilon_{-m} = \epsilon_m$. For every energy level (19) there exists, therefore, another of the form

$$\epsilon = \epsilon_m + 2\mu m H, \tag{19'}$$

which corresponds to the magnetic moment $-\mu m$.

Accordingly, the sum of states $Q = \Sigma \exp(-\epsilon/kT)$ can be divided into two parts

$$Q = Z(H) + Z(-H), \tag{20}$$

the first containing all the terms with exponents of the type (19), the second all those of the type (19'). The two halfs $Z(H)$ and $Z(-H)$ differ only in the sign of H so that the sum of states Q is an even function of this quantity. The average magnetic moment of the system

$$M = kT \frac{d \log Q}{dH} = kT \frac{Z'(H) - Z'(-H)}{Z(H) + Z(-H)} \tag{21}$$

becomes an odd function of H which vanishes for $H = 0$. This fact seems to make this theory inadequate for the explanation of permanent magnetization. We believe, however, that this average is obtained by a type of statistics which has a limited physical reality. In the case of ferromagnetic substances the averaging must be done in a different way. We shall return to this question in section 10 and shall first develop a few equations for the evaluation of the function $Z(H)$. With the help of the formula (12), we can represent each term of the sum of states as $\prod_{j=1}^{r} \exp(-\epsilon_j/kT + 2m\mu H/kT)$, where $2m = N - 2r$. But according to (13) or (16), ϵ_j is a function of the parameters k_j, l_j, n_j and, for a given m, there will be as many different terms as there are possible choices of these numbers. The totality of them is $\exp(2m\mu H/kT) Z_r$,

$$Z_r = \sum_{k_j} \prod_{j=1}^{r} \exp(-\epsilon_{k_j}/kT). \tag{22}$$

We use here the subscript k_j as an abbreviation for the triple of parameters k_j, l_j, n_j. As we found in the preceding section, all the ϵ_{k_j} in the sum (12), or all those in our product, must be different. In other words, our problem is subject to Fermi statistics. Our formula can, therefore, be treated in the usual way: If we consider the expression

$$R = \prod_{k_j} [1 + z \exp(-\epsilon_{k_j}/kT)], \tag{23}$$

then, Z_r is obviously the coefficient of the term of the degree r of the expansion of R in powers of z:

$$Z_r = \frac{1}{2\pi i} \int \frac{R}{z^{r+1}} dz, \tag{24}$$

the path of integration being a circuit round the point $z = 0$.

The half $Z(H)$ of the sum of states is now

$$Z(H) = \sum_{r=0}^{N/2} \exp\left[(N - 2r)\mu H/kT\right] Z_r. \tag{25}$$

Strictly speaking, the expressions R are not the same in the different functions Z_r. But within the approximation of putting k_j equal to k_j', etc. (section 4) we can regard them as identical. Substituting (24) in (25) and summing under the sign of the integral, we find

$$Z(H) = \frac{1}{2\pi i} \int \frac{R}{z^{\frac{1}{2}N+1}} \cdot \frac{(ze^{\tau})^{\frac{1}{2}N+1} - 1}{ze^{\tau} - 1}\, dz, \tag{26}$$

where τ is an abbreviation for

$$\tau = 2\mu H/kT. \tag{27}$$

Although we spoke of one electron in every atom our equations apply as well to the case of several valency electrons of different energy. In fact, each of them will be in resonance only with its own kind in the other atoms and will satisfy a separate secular system. The total sum of states is then the product of the sums of state (26). If there are in an atom two or more electrons in nearly equivalent orbital states and with unlinked or weakly linked spins, the conditions are somewhat different; they are being now investigated by Mr. D. Weinstein. It is probable that our expressions will continue to be true, also in this case, with a change of the physical meaning of the symbol J_1 which will now include both the external and the inner interchange energies.

6. Partial Evaluation of the Sum of States

We consider the case that ϵ_j/kT is, in the average, large in absolute value compared with the quantity τ. In fact the characteristic feature of our theory is Heisenberg's assumption that there is a strong inner directing force due to the energies ϵ_j. The case of small ϵ_j is that when this assumption is not justified; it is, therefore, uninteresting in this connection. We have two terms under the integral (26), let us focus our attention on the second and evaluate it by the method of the steepest descents. In the usual way, we put the integrand equal to e^{ϕ} and obtain the position of the saddle point $z = A$ from the condition $d\phi/dz = 0$:

$$\sum_{k_j} \frac{A \exp(-\epsilon_j/kT)}{1 + A \exp(-\epsilon_j/kT)} - (\tfrac{1}{2}N + 1) - \frac{A}{A - e^{-\tau}} = 0. \tag{28}$$

The second derivative we denote by

$$q = -\sum_{k_j} \left[\frac{\exp(-\epsilon_j/kT)}{1 + A \exp(-\epsilon_j/kT)}\right]^2 + \frac{\frac{1}{2}N + 1}{A^2} + \frac{1}{(A - e^{-\tau})^2}. \tag{29}$$

The position of the saddle point depends on the sign of the energy ϵ_j or of the interchange integral J_1. If the sign is *negative*, A is small compared

with $\exp(-\tau)$: In fact, the left side of the Eq. (28) is negative when $A=0$, but since the exponentials are large, most terms of the sum are, practically, equal to one already for moderate values of A. The sum becomes equal to N and the expression positive. There is a root in between, which does not lie close to the point $z=e^{-\tau}$. Therefore, the last term in (28) is negligible compared with the first two and must be omitted. Neglecting, at the same time, 1 beside $\frac{1}{2}N$, we write it in the form

$$\sum_{kj} \frac{A \exp(-\epsilon_j/kT)}{1 + A(-\epsilon_j/kT)} = \frac{N}{2} \cdot \tag{28'}$$

Turning to the first term of the integrand of (26), we see that it is regular within the path of integration. It has no pole in $z=0$ and the point $z=e^{-\tau}$ lies outside the circuit of integration. The second term represents, therefore, the whole integral which takes the form

$$Z(H) = \frac{1}{(2\pi q)^{1/2}} \prod_{kj} [1 + A \exp(-\epsilon_j/kT)] A^{-N/2} \cdot (1 - Ae^{\tau})^{-1}. \tag{30}$$

It must be noted that A does not depend on the magnetic field because the term containing τ is neglected in the Eq. (28') $Z(H)$ depends on H, only inasmuch as it enters into the last factor of the expression.

The results are quite different in the case of a *positive* energy ϵ_j. A becomes now much larger than $e^{-\tau}$. In fact, for $A=\infty$ the left side of the Eq. (28') is positive. But for comparatively large values of A the terms of the sum become negligible so that the other terms dominate and make the expression negative. The integral retains exactly the same form (30) but it does not represent now the whole function $Z(H)$. The circuit of integration is widened to pass through the distant point $z=A$ and the point $z=e^{-\tau}$ is now within this circuit. Consequently, the first term of the integrand (26) has now a pole in this point and the integral does not vanish but has the value $Z^{(1)}(H)$:

$$Z(H) = Z^{(1)}(H) + Z^{(2)}(H),$$

where

$$Z^{(1)}(H) = \exp \frac{N\tau}{2} \cdot \prod_{kj} [1 + \exp(-\tau - \epsilon_j/kT)] \tag{31}$$

and $Z^{(2)}$ is identical with the expression (30). It is obvious , for physical and for mathematical reasons, that at low temperatures $Z^{(2)}$ is negligible compared with $Z^{(1)}$. In fact, apart from the unimportant factor $(1-Ae^{\tau})^{-1}$, $Z^{(2)}(H)$ is identical with the expression $Z_{N/e}$ of the last section which represents the totality of all terms of the sum of states corresponding to $m=0$. It follows from our expression of the energy (12) that the exponents of these terms are very large, because all the terms of the sum (12) must be different and, consequently, only few of them can have very small values. For this reason $Z_{N/2}$ is negligible, if J_1/T is sufficiently large.

7. Evaluation Continued

We now substitute into our general expressions the formula (16) for the energy term. It will be sufficient to treat the case of the three dimensional lattice as the other cases have no physical interest. We use the approximation $k_j, l_j, n_j = 1, 2, \cdots G$ which is good enough as was shown in section 4. If we define

$$\xi = \frac{\pi k_j}{2G}, \qquad \eta = \frac{\pi l_j}{2G}, \qquad \zeta = \frac{\pi n_j}{2G}, \tag{32}$$

the summation can be replaced by an integration with respect to ξ, η, ζ between the limits 0 and $\pi/2$. Moreover, we use the abbreviations

$$\beta = 2J_1/kT, \qquad \cos 2\xi + \cos 2\eta + \cos 2\zeta = \rho. \tag{33}$$

The terms of the formula (16) become then

$$\epsilon_j = 2\beta kT(\sin^2 \xi + \sin^2 \eta + \sin^2 \zeta) = \beta kT(3 - \rho). \tag{34}$$

Turning to the condition (29′) we can include the first term $3\beta kT$ into the parameter A writing $A' = A \exp(-3\beta)$. We obtain, then,

$$\frac{8}{\pi^3} \int\int_0^{\pi/2} \int \frac{A' \exp \beta\rho}{1 + A' \exp \beta\rho} d\xi d\eta d\zeta = \tfrac{1}{2}. \tag{35}$$

The constant A' can be determined from this relation in the following way. Let us subtract both sides of the equation from 1: The right side remains equal to $\frac{1}{2}$ while the left side has an integrand with the same denominator as (35) but with the numerator 1. The equation can, therefore, be represented as

$$\frac{8}{\pi^3} \int\int_0^{\pi/2} \int \frac{(1/A') \exp(-\beta\rho)}{1 + (1/A') \exp(-\beta\rho)} d\xi d\eta d\zeta = \tfrac{1}{2}.$$

We substitute now $\xi = \frac{1}{2}\pi - \xi'$ and similarly for η, ζ, so that $\rho = -(\cos 2\xi' + \cos 2\eta' + \cos 2\zeta') = -\rho'$, obtaining an equation which has quite the same form as (35) except that A' is replaced by $1/A'$. We conclude from this that $A' = 1/A'$ or

$$A' = 1, \qquad A = \exp 3\beta. \tag{36}$$

If we omit in (30) the unimportant factor $1/(2\pi q)^{\frac{1}{2}}$, $Z(H)$ takes in the case of negative interchange integrals $J_1 < \theta$ the form

$$\log Z(H) = -(3/2)\beta N + \frac{8N}{\pi^3} \cdot \int\int_0^{\pi/2} \int \log [1 + \exp \beta (\cos 2\xi + \cos 2\eta + \cos 2\zeta)] d\xi d\eta d\zeta - \log [1 - \exp (3\beta/2 + \tau)]. \tag{37}$$

Equally simple are the expressions in the case of a positive interchange integral ($J_1 > 0$). Substituting the notations (32) into the expression (31) we find

$$\log Z^{(1)}(H) = \tfrac{1}{2}N\tau + \frac{8N}{\pi^2} \cdot \iint_0^{\pi/2} \int \log\left[1 + \exp\left(-\tau - 2\beta(\sin^2\xi + \sin^2\eta + \sin^2\zeta)\right)\right] d\xi d\eta d\zeta, \tag{38}$$

to this must be added negative term $Z^{(2)}$ of the form (37) which is negligible for low temperatures.

8. Electronic Specific Heats

If the crystal is not subject to magnetic forces ($H=\tau=0$), the two halfs of the sum of state Q become identical and $Q=2Z(0)$. As the free energy is connected with Q by the relation $\psi=-kT\log Q$, it is easy to find the expressions for the internal energy and the specific heat.

We shall see in the next section that the case $J_1<0$ is that of non-magnetic (diamagnetic) bodies. It we denote, in the usual way, $kN=R$ and, moreover,

$$\beta = -\theta/T, \qquad \theta = -2J_1/k, \tag{39}$$

and the number of valency electrons in an atom by s, we find for such materials the temperature dependent part of the internal energy

$$U = sR\theta\frac{8}{\pi^3}\iint_0^{\pi/2}\int \frac{\rho d\xi d\eta d\zeta}{1+\exp(-\beta\rho)}. \tag{40}$$

This expression is but slightly different from Bloch's and agrees with it in every important respect giving results analogous to those obtained by Sommerfeld for free electrons. For high temperatures we can expand with respect to $\beta\rho$ taking only terms of zero and first order. This gives

$$\log Z(0) = Ns(\log 2 - 3\beta/2 + 3\beta^2/16), \tag{41}$$

the temperature dependent part of the internal energy

$$U = -3sR\theta^2/8T,$$

and the specific heat

$$c = dU/dT = 3sR(\theta/T)^2/8. \tag{42}$$

The evaluation for low temperatures can be carried out in a similar way as with the integrals of Fermi and Sommerfeld. The zero order approximation is obtained by taking the integral $\iiint\rho d\xi d\eta d\zeta$ over the part of the cube $0\leqq\xi\leqq\pi/2$, etc., limited by the surface $\rho=\theta$. This term which, in every case, constant and of little interest happens to vanish. For the next approximation it is well to remember that the normal distance between neighboring surfaces $\rho=\cos 2\xi+\cos 2\eta+\cos 2\zeta=\text{const.}$ is equal to $[(\partial\rho/\partial\xi)^2+(\partial\rho/\partial\eta)^2+(\partial\rho/\partial\zeta)^2]^{-\frac{1}{2}}\Delta\rho=\frac{1}{2}[\sin^2 2\xi+\sin^2 2\eta+\sin^2 2\zeta]^{-\frac{1}{2}}\Delta\rho$. Near $\rho=0$ it can be expressed by $\frac{1}{2}[3-2(\cos 2\eta+\cos 2\zeta)^2+2\cos 2\eta\cos 2\zeta]^{-\frac{1}{2}}\Delta\rho$. The factor of $\Delta\rho$ varies between a minimum of $1/2\cdot 3^{\frac{1}{2}}$ and a maximum of $\frac{1}{2}$. As the lower values have a larger weight, we take as a rough average $\frac{2}{3}$ of

the minimum plus $\frac{1}{3}$ of the maximum which is very close to $\frac{1}{4}$. On the other hand, we must know the area of that part of the surface $\rho=0$ which lies inside the fundamental cube. This surface goes through the middle of six edges of the cube which lie at the corners of a plane regular hexagon and also through the center of the hexagon. The surface itself is not quite plane and its area is, therefore, slightly larger than that of the hexagon. We take as its rough approximation the area of the circle circumscribed about the hexagon which is equal to $\pi^3/8$. Multiplying the area by the thickness we find that in the neighborhood of the element of volume element takes the form $d\tau=\pi^3 d\rho/32$ and the correction of the integral becomes

$$\tfrac{1}{2}\int_0^\infty \frac{\rho d\rho}{1+e^{-\beta\rho}} = \frac{\pi^2}{24\beta^2}\,. \tag{43}$$

The internal energy for low temperatures is, therefore,

$$U = (\pi^2/24)sRT^2/\theta, \tag{44}$$

and the specific heat

$$c = (\pi^2/12)sRT/\theta. \tag{45}$$

The expressions are different in the case of magnetic materials ($J_1>0$) and they are derived here for the first time. We have to use the expression (38) with $\tau=O$ and find

$$U = sRT\cdot\frac{8}{\pi^3}\int\int_0^{\pi/2}\int \frac{2\beta(\sin^2\xi+\sin^2\eta+\sin^2\zeta)d\xi d\eta d\zeta}{1+\exp 2\beta(\sin^2\xi+\sin^2\eta+\sin^2\zeta)}\,. \tag{46}$$

For high temperatures ($\beta\ll 1$) we can again expand with respect to the exponent and obtain

$$\log Z^{(1)}(0) = Ns\left(\log 2 - 3\beta/2 + \frac{21}{16}\beta^2\right), \tag{47}$$

while $-Z^{(2)}(0)$ is given by the expression (41). We see that for high temperatures $Z^{(2)}$ is not negligible compared with $Z^{(1)}$ but becomes of a similar order of magnitude, especially, as the factor $1/(2\pi q)^{\frac{1}{2}}$ is no longer small in this case. We introduce the slightly changed definitions

$$\beta = \theta/T, \qquad \theta = 2J_1/k, \tag{48}$$

If we determine U and c for $T\gg\theta$ from $Z^{(1)}$ alone,

$$U = -(21/8)sR\theta^2/T, \qquad c = (21/8)sR(\theta/T)^2, \tag{49}$$

we get the correct dependence on the temperature but coefficients which are slightly too large.

The evaluation for low temperatures $T\ll\theta$ is facilitated by the fact that the exponent retains its positive sign in the whole interval of integration. If β is large, the integrand is appreciable only for small values of the variables ξ, η, ζ. It is permissible to replace the signs by their arguments which lead to

$$U = \frac{3}{2}\frac{1}{(2\pi\beta)^{3/2}}\left[1 - \frac{1}{2^{5/2}} + \frac{1}{3^{5/2}} \cdots\right] sRT,$$

or

$$U = 0.0830 sRT^{5/2}/\theta^{3/2}, \tag{50}$$

and

$$c = 0.208 sR(T/\theta)^{3/2}. \tag{51}$$

The coefficient in this formula is rigorous and not approximate as in (45).

The measurements of Eucken and Werth[7] on the specific heats of iron and nickel show for temperatures between 16° and 22°K considerable deviations from the third power law in the sense of an increase of heat capacity. Eucken himself points out that his observations did not go to sufficiently deep temperatures to establish the law which this excess follows. All we can say is this: If the deviation is attributable to the cause discussed in the preceding paragraphs, its order of magnitude leads to a value of θ of a few hundred degrees and a value of J_1 of a few one hundredths of a volt. If the formula (51), for magnetic materials, were confirmed, it could give us very valuable information. The knowledge of the characteristic temperature θ would be very helpful in elucidating the mechanism of ferromagnetism (compare the next two sections).

It should be pointed out that our theory neglects the possible variability of J_1 and θ with temperature. This is not a rigorous procedure because J_1 must depend to some extent on the distance between the atoms. We presume, however, that this source of error is very small at low temperatures because of the insignificance of the thermal expansion in this region. Neither is the energy of magnetization contained in our expressions but this cause makes an appreciable contribution only in the vicinity of the Curie point.

9. Magnetization

We turn now to the behavior of our crystals in a magnetic field. The magnetic moment is obtained by differentiating the logarithm of the sum of states Q with respect to H

$$M = kT\frac{d \log Q}{dH}. \tag{52}$$

Let us first discuss materials with negative interchange integrals ($J_1<0$, $\beta<0$) for which we have to apply the expression (37). The dependence of $Z(H)$ on the strength of the field is all contained in the last factor $[1-\exp(-3\beta/2+\tau)]^{-1}$ which is very small compared with the other factors. In fact, it does not contain the number of atoms N and this makes its logarithm entirely negligible beside the logarithm of the rest of the expression. Substances with a negative J_1 are, therefore, magnetically inert as far as the electronic interchanges are concerned. We must remember, of course, that our formulas are valid only when β is considerably larger than τ, or $-J_1$ larger than $2\mu H$. For fields of about 10,000 gauss this would mean that $-J_1$ must

[7] A. Eucken and H. Werth, Zeits. anorg. Chemie **188,** 152 (1930).

be larger than 10^{-4} volt. We have pointed out in section 6 that the assumption of very small values for J_1 has little interest as it corresponds to the case when Heisenberg's hypothesis does not apply and the older theory of paramagnetism enters into its rights. The fact that materials with negative interchange integrals of appreciable absolute value must be nonmagnetic is evident directly from the energy expressions (12), (16). The state of vanishing magnetic moment ($m=0$) is, then, that of the lowest energy which dominates, in an overwhelming way, over all other states. We arrive, in this way to the following classification of crystals with respect to their interchange integrals J_1. Materials with large negative values of J_1 are nonmagnetic as far as electronic interchanges are concerned. Materials with small values of J_1 (negative or positive) are paramagnetic. Those with large positive values of J_1 are, potentially, ferromagnetic (compare with next section).

In the next section we shall give reasons for the view that the formula (52) should be applied with restrictions to the case of magnetic materials ($J_1>0$). Here we wish only to investigate what sort of a magnetization curve would result if this equation were always valid. We know from formula (21) that this curve is antisymmetrical of the general character of those pertaining to paramagnetic bodies. What is its slope in the point $H=0$? It is easy to see that in this point the slope is given by the expression

$$\left(\frac{dM}{dH}\right)_0 = kT\left[\frac{d^2 \log Z(H)}{dH^2} + \left(\frac{d \log Z(H)}{dH}\right)^2\right]_{H=0}.$$

We shall be satisfied with low temperatures when β becomes large. The Eq. (38) gives us then the approximation

$$\log Z(H) = \tfrac{1}{2}Ns\tau + Ns\left(\frac{T}{2\pi\theta}\right)^{3/2}\left[e^{-\tau} - \frac{e^{-2\tau}}{2^{5/2}} + \frac{e^{-3\tau}}{3^{5/2}} \cdots\right], \qquad (53)$$

$$(dM/dH)_0 = (Ns\mu)^2/kT, \qquad (54)$$

neglecting a term of the order N. If we refer M to one/cm³ of the material, the expression must be multiplied by n/N (n being the number of atoms per/cm³) and becomes $nNs^2\mu^2/kT$. The values of the constants are as follows $\mu=0.9\cdot10^{-20}$, r is of the order 10^{23}, kT is about $4\cdot10^{-14}$ for room temperature. $(dM/dH)_0$ is, therefore, of the order of magnitude $2s^2N\cdot10^{-4}$. This represents a steep rise of the curve even for very small crystals.

10. Explanation of Primary Ferromagnetism

We shall see in the next section that, in a sense, our Eqs. (52) and (54) describe the magnetization curve of certain systems. However, their physical reality is derived in an indirect way: We believe that the method of the sum of states is inapplicable to a coherent crystal in the conditions which prevail at low temperatures. In fact, what is the physical meaning of the sum of states $Q=\Sigma \exp(-\epsilon/kT)$? It is derived from Gibbs' concept of the ensemble of systems which is equivalent to the succession, *in time*, of the different

states of one given physical system. The mean value of any quantity derived by the use of the sum Q represents, therefore, the *time average* of this quantity. In the cases in which this method has been heretofore used this time average is what we really observe. In the theory of gases states which occur with an appreciable probability deviate but little from the mean state. The time average is very close to the most probable state of the system and very close to, or identical with, the readings of our instruments which always integrate over a finite, though sometimes small, time. The situation is entirely different in our theory of magnetization. It is true that positive and negative magnetic moments will occur equally often in the course of long periods and the time average will vanish. But this time average is not at all close to the most probable state. On the contrary, under the conditions mentioned, it coincides with one that is so improbable that it practically never occurs. To show this, let us compute the time average of the *absolute value* of the magnetic moment in the absence of a field.

$$\overline{|2\mu m|} = kT(d \log Z(H)/dH)_0 = Ns\mu[1 - 0.058(T/\theta)^{3/2}]. \tag{55}$$

This means that, on the average, the crystal is in a state close to magnetic saturation. When we further compute the relative mean quadratic deviation from this time average

$$\overline{(|m| - \overline{|m|})^2}/\overline{|m|}^2 = (0.37/Ns)(T/\theta)^{3/2}, \tag{56}$$

we find that the spread around the mean value is extremely small. These formulas show us that the crystal is, practically, always in a state approaching saturation. The time average zero is explained by the fact that the two senses of the magnetic moment are equally probable so that the crystal, at irregular intervals, spontaneously changes its magnetic polarity.

These results refer only to the case of low temperatures. When the temperature becomes sufficiently high, the state of vanishing magnetic moment is no longer one of low probability. This is apparent from our expression (41) valid for high temperatures. This expression is, practically, identical with the logarithm of $Z_{N/2}$ of section 6 representing the totality of all the terms of the sum of states pertaining to the magnetic moment $m=0$. The Eq. (41) shows us that this term has a negative exponent and is small as long as $3\beta/2-3\beta^2/16>\log 2$. But when the temperature rises the inequality is reversed into $3\beta/2-3\beta^2/16<\log 2$, the exponent becomes positive, and the term large. It is true that, in this form, the argument is not quite convincing: The formula is derived for small values of β and it is not certain that it still holds when $3\beta/2-3\beta^2/16$ becomes equal to log 2 or smaller. However, we get qualitatively the same result when we go back to the original definition $Z_{N/2}=\Sigma \exp(-\epsilon/kT)$ the summation being extended over all possible energy levels compatible with the condition $M=0$ or $r=N/2$. Now, let us look at our energy expressions (12) and (16). The number of terms in the sum (12) depends on the magnetic moment $2m$; it has its minimum 0 for saturation ($2m=N$) and its maximum $N/2$ for the nonmagnetic state ($2m=0$). As all the terms of the sum must be different, the mean energy also increases from

$\epsilon = 0$, for saturation, to a huge value, for the conditions of vanishing magnetic moment. It is easy to see that the energy excess of the state $2m = 0$ over the state $2m = N$ cannot be smaller than $1.45NJ_1$ (or $0.72kN\theta$) and not larger than $2.42NJ_1$ (or $1.21kN\theta$). On the other hand, the number of different energy levels pertaining to the state $2m = 0$ is $N!/(N/2)!(N/2)!$ or, within the accuracy of Stirling's formula 2^N. We find, therefore, for $Z_{N/2}$ the limits

$$N(\log 2 - 1.21\theta/T) < \log Z_{N/2} < N(\log 2 - 0.72\theta/T).$$

Owing to the large factor N, the change of $Z_{N/2}$ with T is very rapid. In the region $\log 2 > 1.21\theta/T$;(or $T > 1.8\theta$) it becomes a very large number; in the region $\log 2 < 0.72\theta/T$ (or $T < 1.04\theta$) it is a very small number. The transition from a high probability to a low probability of the unmagnetized state occurs somewhere between these two limits at a point which is roughly of the order

$$T_0 \approx 1.5\theta. \tag{57}$$

The conditions are the same for the other states of magnetization of our crystal ($2m = N - 2r$). As r increases from 0 to $N/2$ the individual terms of the groups Z_r defined by the Eq. (22) become smaller and smaller because their exponents grow in absolute value. On the other hand, the number of terms increases with r being equal to $N!/(N-r)!r!$ For high temperatures this second effect overweighs the first: The probability of a state increases with $r = \frac{1}{2}N - |m|$ and has its maximum for the unmagnetized state $m = 0$. The characteristic temperature above which such a distribution exists is, nearly, independent of r and approximately given by the Eq. (57). For temperatures $T > T_0$, the most probable state and the time average derived from the sum of states coincide and the sum of states has a good physical meaning. The crystal will be unmagnetized or very little magnetized. The opposite is true for low temperatures. The negative exponents are so large that all Z_1 are negligible but those pertaining to very small r. We have here the conditions analysed in the beginning of this section. *The point T_0 of Eq. (57) is, therefore, the primary Curie point of a coherent crystal*: Below it the crystal is a state of spontaneous magnetic saturation, above it, in a, practically, unmagnetic state.

Let us now return to the discussion of conditions below the Curie point. As we have seen, a coherent crystal is, then, in a state very near to complete magnetic alignment but its magnetic polarity spontaneously reverses itself at irregular intervals. What is the length of these intervals? Any change of magnetization is due to perturbations coming from the irregular thermokinetic state of the medium around the crystal and of the crystal itself. If there were no mutual magnetic interactions between the electronic spins, every electron would respond to these perturbations independently. In order to change its polarity, the crystal would have to pass through the extremely improbable state of vanishing magnetic moment and such an event could occur only once in an eon. In reality, there exists some measure of magnetic linking between the spins and they can turn over in groups so that the crystal

need not pass through the most improbable state. However, if the crystal is not small in size, the periods involved are still very long. The estimate can be made in a way analogous to that used by Bloch,[8] in a somewhat different case, from the magnetic linking of the spins and from the moment of momentum which the crystal has to pick up in order to reverse its magnetization. We call a crystal *coherent* if Slater's equation applies to it as to a unit, the numbers N and G determining its size. If there existed coherent crystals of a considerable size they would be always in a state of complete magnetization in some direction and a field could only change this direction but not demagnetize them. This is contrary to observation and we are compelled to conclude that the coherence extends only over microscopic elements. Even a well developed monocrystal must have a *block structure*. Slater's equation and all the consequences contained in it which we have developed in the preceding sections applies only to the separate blocks. It is unimportant for us whether these discontinuities of a crystal are irregular (in the sense of the *mosaic structure* surmized by Darwin and Smekal) or form a regular *secondary structure*, as advocated by Zwicky. The size of a few hundred atoms in each direction G, postulated by these authors, would insure a sufficiently frequent spontaneous change of magnetization.

We arrive, in this way, at the following picture of the constitution of ferromagnetic materials. They are built up of microscopical blocks which are in a state of permanent spontaneous magnetization. The polarity of this magnetization is subject to frequent spontaneous changes. As far as our simplified theory goes the electronic spins might have any direction and the magnetic axis could rotate freely within the block. In reality there must exist certain directions of preference related to the main crystallographic directions because of the interactions between spin and orbital momentum.[9] The magnetic polarity will alternate between these preferred directions. The permanent magnetization of these micro-crystalline blocks is what we designate as the *primary ferrogmagnetism*. The *secondary magnetization* of the bulk of a ferromagnetic substance is produced by the alignment of the magnetic axes of the primary elements. While this picture is, by no means, new we claim to have demonstrated, for the first time, that the permanent magnetization of the blocks is not a hypothesis but a necessary consequence of the exclusion principle.

11. Remarks on Secondary Magnetization

Already in Maxwell's *Treatise* we find the opinion that ferromagnetic materials are built up of permanently magnetized microcrystalline blocks. He believed that a magnetic field produces an alignment of the magnetic axes by rotating the whole block. Although this seemed little probable because of the enormous viscosity of these materials, I thought Maxwell's hypothesis sufficiently interesting to have it tested experimentally. A few years ago I caused Mr. Yensen[10] to look for a possible rotation using the Hull-Debye x-ray

[8] F. Bloch, Zeits. f. Physik **74,** 295 (1932).

[9] Compare F. Bloch and G. Gentile, Zeits. f. Physik **70,** 395 (1931).

[10] T. D. Yensen, Phys. Rev. **32,** 114 (1928).

method. It was not very surprising that the result was negative. According to our results of the preceding sections, Maxwell's assumption is unnecessary because the magnetic polarity of a block alternates while the block itself remains in place. In modern times the idea of permanently magnetized microcrystals was used with great success by Weiss[11] in the case of polycrystalline materials. Recently, Bitter[12] gave a considerable amount of experimental evidence which tends to show that it applies also to monocrystals.

The objective of this paper and its main result is to establish the primary ferromagnetism of the microcrystalline blocks. There exists a considerable literature about secondary magnetization and we limit ourselves, in this connection, to a few remarks. The agglomeration of blocks represents a large number of similar systems side by side. We can interpret, therefore, the statistical averages of section 9 in a new way. They represent as well the time averages for a single block as the actual mean values for the agglomerate. The simplifying assumption is implied that the blocks are all equal and that the magnetization has one of two opposite directions. It would be easy to drop these restrictions but hardly worth our while. On the contrary, we simplify still further by dropping the small second term of (53), i.e., by assuming that our blocks are completely magnetized. Eq. (52) gives then for the magnetization (per unit volume of the agglomerate)

$$M = ns\mu \operatorname{tgh} \frac{Ns\mu H}{kT} .$$

There are two effects left out of account in this formula, which can modify it sufficiently to produce a spontaneous and permanent secondary magnetization. The first is the field produced at a given point by the rest of the system. According to Lorentz,[13] this cause can be taken care of by substituting instead of the outer field H the actual strength of field $H' = H + \alpha M$. In Lorentz' cases α was of the order $\frac{1}{3}$ but it is hard to say what it is equal to in our case. The condition for secondary permanent magnetization would then be $(dM/dH)_0 < 0$ or $nNs^2\mu^2/kT > 1/\alpha$. The fact that monocrystals of iron and cobalt have little or no remanence indicates that this cause is not quite sufficient to produce permanent magnetization and points to a rather small size of the blocks or to a small value of α. The other effect which distorts the magnetization curve is the influence of mechanical stresses put forward by Akulov[14] and discussed in detail by Gans[15] and by Becker.[16] These two causes seem entirely sufficient to account for all the phenomena of secondary ferromagnetism. We wish, only, to point out that they may give rise to a secondary Curie point which would be observable by magnetic methods if it happened to be lower than the primary Curie point determined by the Eq.

[11] P. Weiss, Phys. Zeits. **9,** 361 (1908).
[12] F. Bitter, Phys. Rev. **38,** 528 (1931); **39,** 337 (1932).
[13] H. A. Lorentz, Enzyklopadie der math. Wiss. V 14, section 37.
[14] N. S. Akulov, Zeits. f. Physik **64,** 817 (1930).
[15] R. Gans, Schriften d. Königsberger gelehrten Gesellschaft **8,** 33 (1931).
[16] R. Becker, Zeits. f. Physik **62,** 256 (1930).

(57). The problem of relating these theoretical Curie points to those found by observation is not simple and we reserve it for another communication.

Note added in proof. It was overlooked by the writer that, in another paper (Leipziger Vortraege 1930, p. 67), Bloch gives the proportionality of the specific heat of ferromagnetic materials with $T^{3/2}$. However, the agreement with our results is more or less accidental, as Bloch's treatment is quite different and makes use of the two assumptions, pointed out at the end of our section 4, which we do not consider as permissible.—The experimental test of this law is complicated by the fact that the elements of the iron group contain two kinds of valency electrons, the first kind having, presumably, a negative interchange integral, the second a positive one. (Compare: Slater, note 6). The complete electronic specific heat is, then, to be represented by the sum of our formulas (45) and (51): $c = R[\alpha_1 s_1 T/\theta_1 + \alpha_2 s_2 (T/\theta_2)^{3/2}]$, $\alpha_1 = \pi^2/12$, $\alpha_2 = 0.208$. Only θ_2 is related to the primary Curie point.

ERRATA

The Shape of an X-ray Line

By ARCHER HOYT
Department of Physics, Cornell University

(Phys. Rev. **40**, 477, 1932)

It has been called to the author's attention that there are two errors in Table I of the article by the above title.

Spencer[1] has been misquoted. The values published for half-maximum *half-widths* are in reality half-maximum *full-widths*. The values for Mo $K\alpha_1$ and Cu $K\alpha_1$ should be reduced by a factor $\frac{1}{2}$. The values which *should* appear in the table of half-widths at half-maximum ordinate are

for Mo $K\alpha_1$	4.8 sec.	or 0.14 X.U.
for Cu $K\alpha_1$	10.8 sec.	or 0.305 X.U.

[1] Spencer Phys. Rev. **38**, 630 (1931).

LETTERS TO THE EDITOR

Prompt publication of brief reports of important discoveries in physics may be secured by addressing them to this department. Closing dates for this department are, for the first issue of the month, the twenty-eighth of the preceding month; for the second issue, the thirteenth of the month. The Board of Editors does not hold itself responsible for the opinions expressed by the correspondents.

The Vapour Pressure of Sulphur at 50°C

Earlier workers[1] have obtained widely differing results for the vapour pressure of sulphur at 50° *viz.*, 0.00008 and 0.00036 mm of mercury. We made several determinations using a new method.

The apparatus is shown in Fig. 1. *A* is a drying tower containing soda-lime and calcium chloride. *B* is a thermostat heated and controlled electrically and maintained at 50°C. *C* is a glass spiral containing sulphur (purified by distillation in vacuo) while *D* is a Pyrex tube heated electrically to 300°C. The water jacket *E* cools the issuing gases which are then absorbed in *N*/100 iodine solution in wash bottle *F*. Bottles *G*, *G* contain potassium iodide solution to trap any iodine carried over.

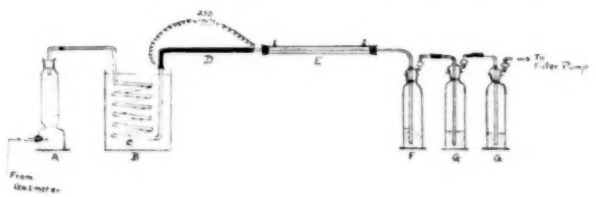

Fig. 1.

The procedure is as follows. Air is drawn through the apparatus by means of a filter pump fitted with constant leak. The air passes first through a gas meter and drying tower and then over the sulphur in the spiral which is at 50°C. The sulphur is oxidised to sulphur dioxide in the Pyrex furnace (owing to the very high partial pressure of oxygen this takes place almost quantitatively) and the issuing sulphur dioxide after cooling is absorbed in the *N*/100 iodine solution. After passing air for some time (24–48 hours) at a rate of about 10 litres/hour the wash bottles are detached and the iodine in them titrated with a sodium thiosulphate solution previously standardized against the iodine.

Thus we find the loss in iodine and hence the weight of sulphur dioxide formed. Knowing from the gas meter the volume of air which has passed and also the barometric pressure and the temperature at the meter, we can readily calculate the vapour pressure of sulphur at 50°C.

We obtained as the mean value *0.000155 mm Hg* assuming all the sulphur present as S_8.

In conclusion we wish to offer our best thanks to Mr. A. Ritchie, B.Sc for his invaluable help and advice.

R. R. H. Brown
J. J. Muir

Edinburgh University,
Chemistry Department,
April, 1932.

[1] Ruff and Graf, Z. Anorg. Chem. **58** 210 (1908); Greuner, J.A.C.S. 1399 (1907); West and Menzirs., J. Phys. Chem. **33**, 1887 (1929).

Variation of the Cosmic Rays with Latitude

Definite differences in the intensity of the cosmic rays at different latitudes are shown by our measurements, which have ranged from 47° north to 46° south. As far as they have gone, these measurements indicate a uniform variation with latitude, showing a

minimum at or near the equator, and increasing intensity toward the north and south poles. At sea level, the difference between the intensity at latitude 45° and 0° is roughly 16 percent, whereas at an elevation of 9000 feet the difference is about 23 percent. This would indicate that it is the least penetrating part of the cosmic rays which varies most rapidly with latitude. No significant variations with longitude have been noted.

Our present experiments are being made with a 10 cm steel sphere filled with argon at 30 atmospheres. The ionization current is measured by a Lindemann electrometer, the readings being made in such a way that the effect of any insulation leak is eliminated. To shield from the effects of the local radiation, a spherical shell of 2.5 cm copper and two spherical shells each of 2.5 cm lead, making a total of 7.5 cm of lead and copper, are used. The effect of the local gamma-rays traversing this shield is determined by comparing the ionization with and without the outer shell of 2.5 cm lead. All the readings are referred to the ionization by the gamma-rays from a capsule of 1.3 mg of radium placed at 1 meter from the center of the ionization chamber as a standard.

TABLE I. *Cosmic ray intensity at different localities*
(Ions per cc per sec. through 5 cm Pb, 2.5 cm Cu and 0.5 cm Fe)

Location	Lat.	Long.	Elev.	Barom.	I_C	I_L	Date
1 Mt. Evans	40°N	106°W	14,200ft	17.61in	6.88 ions	0.57	9/31
2 Summit Lake	40 N	106 W	12,700	18.70	5.84	0.34	9/31
3 Denver	40 N	105 W	5300	24.8	2.93	—	9/31
4 Jungfraujoch	47 N	6 E	11,400	19.70	5.08	0.51	10/31
5 Haleakala	21 N	156 W	9300	21.47	3.35±0.05	0.60	4/32
6 Idlewild	21 N	156 W	4200	25.99	2.40±0.05	0.37	4/32
7 Honolulu	21 N	158 W	70	30.09	1.89±0.02	0.11	4/32
8 S. S. Aorangi	4 S	173 W	60	29.65	1.83±0.05	0.32	4/32
9 Southern Alps	44 S	170 E	6700	23.69	3.39±0.05	0.22	4/32
10 Southern Alps	44 S	170 E	3900	26.10	2.70±0.04	0.21	4/32
11 Dunedin	46 S	170 E	80	30.08	2.16±0.03	0.11	4/32
12 Wellington	41 S	175 E	400	29.85	2.16±0.03	0.12	5/32

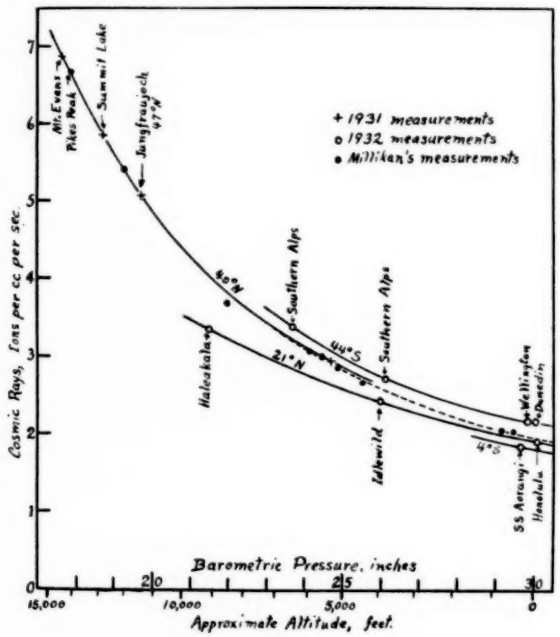

Fig. 1.

In Table I are listed the results of our various measurements. The column marked I_C represents the ionization in air at atmospheric pressure expressed in ions per cc per second due to the cosmic rays traversing our 7.5 cm filter. I_L represents the ionization due to the local rays passing through this filter expressed in the same units. Measurements 1 to 4 were made with different apparatus, using compressed air instead of compressed argon in the chamber, and lead and copper shields of different shape. These results are thus not strictly comparable with the later ones, though the difference is probably not more than 2 percent.

In Fig. 1 these results are shown graphically, and are compared with those of Millikan and his collaborators (Phys. Rev. **37,** 235 1931) as shown in the dotted line. It will be noted that for corresponding barometric pressures our measurements show consistently more intense radiation at the higher latitudes. The high altitude measurements in Hawaii are of especial interest as indicating that the difference with latitude is greater for the less penetrating rays.

Millikan's measurements were made at latitudes varying from 34°N to 59°N, most of them being at about 40°N. Our measurements near this latitude are in satisfactory agreement with his considering the slight differences in experimental arrangement. We do not however confirm the independence of the intensity of latitude which he reports.

The variation of the cosmic rays with atitude which these measurements show is of ust the kind to be expected if the rays consist of electrically charged particles which are deflected by the earth's magnetic field, the less penetrating rays being the more strongly affected.

This letter is the first report of an extensive program involving similar measurements by many physicists in widely distributed parts of the world. The design and construction of the instruments are due chiefly to Professor R. D. Bennett of Massachusetts Institute of Technology, with L. N. Ridenour, Dr. J. A. Hopfield and Dr. E. O. Wollan. Professor J. C. Stearns of the University of Denver with J. A. Longman, L. N. Ridenour and W. Overbeck made the measurements at Denver and Mt. Evans. Dr. Marcel Schein and Dr. Bernhard Frey of the University of Zürich cooperated in the measurements at the Jungfraujoch. Professor Harry Kirkpatrick of the University of Hawaii made the local arrangements in Hawaii, and with Professor W. H. Eller and Mr. I. Miyake assisted in the measurements. In New Zealand, Professor P. W. Burbidge of Auckland University College organized the various expeditions, and with Dr. C. M. Focken of the University of Otago helped with the observations. Valuable assistance has also been given by A. A. Compton and others. Without the cordial cooperation of the various universities where the work has taken us, the measurements would have been much more difficult.

ARTHUR H. COMPTON

University of Chicago,
The Tasman Sea,
May 7, 1932.

On the Production of Ultra-Short Electromagnetic Waves

In a note published in this journal, Mr. G. Potapenko gave an account of his interesting investigations in the field of the ultra-short electromagnetic waves.

The tubes that the author used in his experiments had pure tungsten filaments, as it was impossible to observe ultra-short waves with tubes having other kinds of filaments.[1]

As a result of some experiments I carried out recently, I found that, under particular conditions, oscillations of very high frequency may be obtained by using tubes with a heater element and a cathode coated with oxides.

Different tubes have been tested, namely, tubes of the type: −27 R.C.A. (American); E. 424 Philips (Dutch); R.E.N.S. 1104 Telefunken (German).

Although the common Barkhausen-Kurz arrangement may be used, that of Tank and Schiltneck[2] was found to be more suitable, the only modification being to connect the cathode to the negative terminal of the filament. The heating current must have a value which is only 80 percent − 90 percent of its normal rated value.

Although unnecessary if the latter experimental arrangement is used, a slight positive

[1] Phys. Rev. **39,** 630 (1932).
[2] Helv. Phys. Acta. **1,** 100 (1928).

plate potential (0.5 to 2 volts) increases the amplitude of oscillations. Under these conditions a plate current is present even when the tube is not oscillating; this current being obviously increased when oscillations set up. This increase of current may vary (in accordance with the particular tube, etc.) between 10 and 200 microamperes. As the grid potential runs between 20 and 100 volts, a band of wave-lengths between 180 and 50 cm can be obtained.

AMEDEO GIACOMINI

Istituto di Fisica della Regia Università,
Milano, Italy,
June 1, 1932.

Predissociation in Nitrogen

Predissociation in nitrogen is best illustrated by the first positive bands $B_{13}\rightarrow A_9$, $B_{14}\rightarrow A_{10}$, $B_{15}\rightarrow A_{11}$ and by the bands $B_{13}\rightarrow A_8$, $B_{14}\rightarrow A_9$ and $B_{15}\rightarrow A_{10}$.[1] On small dispersion spectograms most of the bands of the first positive group possess three fairly strong heads, which correspond to transitions from the three components of the $^3\pi$ initial state of these bands to the $^3\Sigma$ state. In the above-mentioned bands however, two of the heads are almost completely missing. About a year ago, I reported the excitation of these missing heads in mixtures of nitrogen and oxygen,[2] in which the concentration of nitrogen was relatively small compared with that of the oxygen. At that time, the explanation of this effect was proposed to be the prevention of the normally occurring predissociation by electric fields, which were introduced during collisions. There is an obvious alternative explanation in which electric fields play no direct role, namely, the effect of increasing the number of molecules in the normally predissociating levels by collisions which give rise to transition from the non-predissociating component of the $^3\pi$ state to the two predissociating ones. The energy differences between the components of the $^3\pi$ state are very small and such transfers should take place without difficulty. The effect of pressure on predissociation has been discussed for AlH by Stenvinkel[3] and others, and this effect in nitrogen is in some respects, similar to the one in AlH. There is a difference, however, for in nitrogen the collision transfers are between components of an electronic state, whereas in AlH they are between rotational levels.

The occurrence of this effect in nitrogen suggests a similar explanation for the enhancement of predissociation in iodine by argon.[4] Mulliken's assignments of $^3\pi_0{}^+$ to the upper state of the visible halogen bands, makes this state a triplet state and hence provides the two other components, $^3\pi_1$ and $^3\pi_2$, which are necessary in order that the iodine problem be similar to the one in N_2. In iodine, the two levels $^3\pi_1$ and $^3\pi_2$ may themselves be unstable or they may combine with a third unstable level as in nitrogen.

Before the appearance of Van Vleck's[5] paper on the theory of the magnetic quenching of iodine fluorescence, I had proposed to myself the problem of studying the effect of magnetic fields on the predissociation in nitrogen, in order to observe the effect of magnetic fields on collisions. It was thought at the time that the magnetic field might be able to enhance collision transitions from the non-predissociating $^3\pi$ state to the two predissociating ones, and thus be able to quench the fluorescence. I am not aware of any theoretical reason for believing that this explanation has any value, but in view of the effect of collisions on predissociation both in iodine and in nitrogen, it seems worth while to suggest that the effect of magnetic fields on collisions may be of some significance.

JOSEPH KAPLAN

University of California
at Los Angeles,
June 3, 1932.

[1] J. Kaplan, Phys. Rev. **38,** 373 (1931).
[2] J. Kaplan, Phys. Rev. **38,** 1079 (1931).
[3] Stenvinkel, Zeits. f. Physik. **62,** 201 (1930).
[4] Turner, Phys. Rev. **38,** 574 (1931).
[5] Van Vleck, Phys. Rev. **40,** 544 (1932).

The Isotopic Weight of H^2

The mass of neutral H^2 has been measured as 2.01353 ± 0.000064 by comparison of the mass of $(H^1H^1H^2)^+$ with He^+ on fourteen spectra obtained with the mass-spectrograph recently described.[1] For the purpose of these measurements the mass of helium was taken as 4.00216, that of hydrogen 1.00778, and for the electron 0.00055 units on the $O^{16}=16$ scale.[2] The equivalent packing fraction of H^2 is 67.6 parts in 10,000. The energy of binding corresponds to 0.00203 mass units or approximately 2×10^6 electron-volts if the nucleus is composed of three entities, two protons and one electron. It is interesting to speculate that if the H^2 nucleus is composed of one proton and one Chadwick neurton[3] of mass 1.0067, then the binding energy would be 9.5×10^5 electron-volts and the H^2 nucleus would be loosely bound compared to He and other light nuclei. Professor Urey has pointed out to the author that for a stable configuration the mass of a free neutron must be greater than 1.00575.

The probable error of a single determination was ±0.00031. The observations were weighted from one to three depending on the definition of the lines of specific spectra and the internal consistency of the mensuration of the position of those lines. The sum of the weight factors was twenty-three.

$(3H^1)^+$ and He^+ provided the dispersion measurements for these spectra. The presence of $(H^1H^2)^+$ can only introduce a negligible correction. Earlier unreported work on the carbon and oxygen hydrides has shown that the mass scale is linear to within one part in ten thousand over the first eight centimeters of the plate, which includes the region used in this investigation.

Lines on the mass-spectrograph plates corresponding to a mass of 4.0285 were attributed to $(H^1H^1H^2)^+$ because; (1) no lines of comparable intensity appeared in this position when ordinary commercial hydrogen was used in the discharge tube; (2) the probability of the formation of $(2H^2)^+$ was much less than the probability of $(H^1H^1H^2)^+$ under the conditions existing in the discharge tube, i.e., $(3H^1)^+$ was present from 100 to 5 percent of $(2H^1)^+$; (3) the mass is less than the mass of $(4H^1)^+$ by an amount outside the limit of error.

Two samples of enriched hydrogen were used. Dr. Bleakney gave the writer a sample of the hydrogen he had examined,[4] and Dr. Brickwedde and Professor Urey[5] sent a sample of hydrogen with a slightly higher concentration of H^2.

The high value of the mass of H^2 makes it very difficult at this time to resolve the two doublets $(3H^1)^+$, $(H^1H^2)^+$, and $(2H^1)^+$, H^{2+}. It may be possible to resolve these doublets when more concentrated samples of hydrogen are available so that the lines of the respective ions can be obtained of comparable intensity. The doublet $(H^1H^1H^2)^+$, $(2H^2)^+$ will be difficult to resolve in any case as the respective masses differ by only one part in 2000.

The author is particularly indebted to Dr. Brickwedde who made this investigation possible by his preparation of the samples of enriched hydrogen at the U. S. Bureau of Standards, to Professor Urey and Dr. Murphy of Columbia University who placed a tested sample at the disposal of the author and to Dr. Bramley and Dr. Bleakney for their interest and helpful discussions.

KENNETH T. BAINBRIDGE

Bartol Research Foundation of the Franklin Institute, Swarthmore, Pa.,
June 21, 1932.

[1] K. T. Bainbridge, Phys. Rev. **40,** 130 (1932).

[2] F. W. Aston, Proc. Roy. Soc. **115,** 487 (1927).

[3] J. Chadwick, Proc. Roy. Soc. **136,** 702 (1932).

[4] W. Bleakney, Phys. Rev. **39,** 536 (1932); Phys. Rev. **41,** July 1, 1932.

[5] H. C. Urey, F. G. Brickwedde, and G. M. Murphy, Phys. Rev. **40,** 1 (1932); **39,** 164 (1932).

BOOK REVIEWS

Photoelectric Phenomena. ARTHUR LLEWELLYN HUGHES AND LEE ALVIN DUBRIDGE. Pp. 531. McGraw-Hill Book Company, Inc., New York, 1932. Price $5.00.

The preparation of an up-to-date text on "Photoelectric Phenomena" constitutes an almost hopeless undertaking. So many and varied are the researches in this field published in current technical journals that to digest, organize and include the subject matter in an orderly discussion would be next to impossible. The task is doubly difficult due to the fact that so little is known concerning the nature of the photoelectric effect and so many theories have been advanced to explain the oftimes conflicting results obtained by faulty experimental technique.

It is for these very same reasons, however, that there has been felt so keenly the need of a treatise coordinating the data so far obtained; and it is a source of considerable satisfaction to investigators in the field of photoelectricity to discover that "Electrical Phenomena" by Hughes and DuBridge so admirably meets these needs. The fact that several rather important papers published within the year seem to have been overlooked in its preparation, particularly in the field of photoconductivity, and the almost reportorial presentation of the data resorted to by the authors with little or no attempt to pass judgment on the relative merits of the research are to be expected rather than criticized in view of what has already been said.

"Photoelectric Phenomena" is a comprehensive well written, carefully edited and attractively published volume, far eclipsing any other volume on the subject published in the English language and perhaps being second to none emanating from the continent.

Defining a "photoelectric effect as one in which radiation causes a separation of electron from atoms so that an electric current may be caused to flow," the authors set out "to give a concise yet comprehensive survey of one of the many fields of physics." And within the fourteen chapters comprising the book they accomplish just that. The text, after a brief introductory chapter, consists of four chapters on the emission of electrons from metals, one chapter on the theories of photoelectric emission, a chapter each on what they have called "volume" photoelectric effects, viz., ionization of gases by light, and photoconductivity, two chapters on the photovoltaic effect and photoelectric effect in liquids and dielectrics, a chapter on the photoelectric effect of x-rays and γ-rays, a chapter each on photoelectric technique and applications and ends with a discussion of miscellaneous phenomena. The appendix includes a convenient table of physical constants, important photoelectric energy relations, a periodic table of the elements and a summary of illuminating engineering units and nomenclature. A very complete running bibliography is included in the form of footnotes with a name index and subject index at the end of the text. The book is printed in easily readable type and the sketches and diagrams are many and clear.

Unlike previous publications in English, too much valuable space is not devoted to a long historical account of the pioneering work or to the multitudinous applications of photocells. The text is essentially up-to-date and stresses important problems of the present. The order of presentation, chapter by chapter, is: first, ample citation of and reference to early related work, secondly, a complete resumé of all researches since that time, and finally a summary and discussion which generally includes a concise outline of the problems still unanswered. After presenting the researches of others with such reportorial excellence, this discussion of the outstanding problems is especially commendable. Incidently, it reflects the thorough understanding of the field possessed by the authors.

One of the most praiseworthy features of the presentation is the differentiation between "spectral selectivity" and "polarization selectivity." Yet, after keeping these phenomena separated, as they should be, the authors, without justification perhaps, advance the doubtful suggestion that the latter is always associated with the former. Unfortunately, at the very beginning of this chapter a confusing error crept into the definition of polarization selectivity.

There are undoubtedly those who will criticize the order of presentation in the chapter on

the selective photoelectric effect and that on photoconductivity, just as they will the arbitrary classification of photoelectric phenomena as "surface photoelectric effects," "volume photoelectric effects" and "photovoltaic effects." This might follow if the reader is accustomed to the style of Gudden and Pohl in their volume bearing the same title as the one under discussion. In and of itself, however, the subject matter in "Photoelectric Phenomena," by Hughes and DuBridge will be found to be wisely classified, and coherently, competently and clearly presented. Even the chapter on "Theories of the Photoelectric Effect," though brief and undertaken apologetically by the authors, is entirely adequate for such a volume. Moreover, the discussion on photoelectric technique is pertinent and practicable, and can well be accepted as reflecting the experience of careful research workers in the field of photoelectricity.

A. R. OLPIN
Bell Telephone Laboratories, Inc.,
New York, N. Y.

Cartesian Tensors. HAROLD JEFFREYS. Pp. 93, The Macmillan Company, New York 1931. Price—$1.75.

According to the author's preface this little book is intended as a frank attempt to popularize the tensor methods for systems of rectangular axes, even to the extent of eliminating the usual extension development of vector analysis. Even within the realm of classical physics, to say nothing of relativity, one can find good reason for sympathizing with this plan. As examples, one thinks of the inertial ellipsoid of a rigid body and the polarization ellipsoid of a molecule. In these, and in other similar instances, the tensor notation seems much preferable to the Gibbsian dyadics.

The main emphasis in the book is on applications of tensors in classical physics. The fact that only twenty short pages are needed for the development of the tensor calculus itself should recommend it. The applications are made to dynamics of particles and rigid bodies, force systems, electrodynamics (potential theory), elasticity, and hydrodynamics. The discussions of the individual topics are too condensed to give more than a taste but taken together they give the student a valuable summary. The explanations are stated and easily followed.

E. L. HILL
University of Minnesota

Eenige Onderzoekingen over Paramagnetisme. E. C. WIERSMA. Pp. 122. **Paramagnetische Eigenschaften von Salzen.** C. J. GORTER. Pp. 112.

It is perhaps well to call attention in these columns to these two interesting Leiden dissertations devoted to topics in paramagnetism. Both dissertations are primarily the outgrowth of original experimental investigations, but at the same time enter into theoretical considerations with considerable detail. It is to be particularly commended that attention is focused on the newer quantum-mechanical developments rather than on obsolete classical models, as is too often done. In fact, over half of Wiersma's monograph is devoted to theory. The remainder is devoted to his experimental methods and results. The latter include the behavior of O_2 and NO at low temperatures, which is of importance in connection with theory for gases, and some interesting cryomagnetic anomalies in solid iron chloride, in which the susceptibility depends upon field strength in a peculiar way. Gorter's monograph is particularly valuable for the care with which it documents the existing determinations of susceptibilities and magneton numbers in the rare earth and iron groups. It is probably the best available compendium of such data. Gorter's own measurements have been indispensable in this field. They have corrected anomalies reported by other observers for hydrated nickel sulphate. Also his observations on Pr and Nd salts at low temperatures have enabled theoretical physists to estimate quantitatively the crystalline field.

J. H. VAN VLECK
University of Wisconsin

"Einführung in die Mechanik und Akustik" von R. W. Pohl, Zweite Auflage, Pg. 251 +viii, Figs. 440. Julius Springer, 1931. Price RM 15.80.

This volume constitutes an important contribution to the presentation and the elucidation of the introductory principles of mechanics. It is well organized and the author's style of a simple, plain unfolding of the facts of nature as revealed by experiment is a very pleasing one for the reader. The illustrations are numerous, 440 altogether, and their quality of reproduction is of a very high order. The device of silhouetting figures is frequently used and with good effect, and the experiments described exhibit methods of highly refined technique.

The first eight chapters, consisting of 115 pages, treat the mechanics of particles and of rigid bodies up to and including rotational motion and accelerated reference systems. The next two chapters of 54 pages deal with the mechanics of fluids, and these complete the mechanics division of the book. The division on acoustics consists of 75 pages and is divided into two chapters, the first of which deals with vibration theory and the second with the theory of wave motion and radiation.

As a groundwork for mechanics force is taken as a fundamental idea, and in pleasing contrast with so many texts which adopt this method the definition of force measurement is given at the outset. It is true that the author introduces the term mass simultaneously with that of force and also a convention for its measurement. Newton's second law is then an empirical relationship between force and mass as defined.

This mode of development of fundamental ideas is a possible one and is one that is very commonly used. It has certain advantages and its desirability is a matter for personal opinion. On the other hand the author expresses views to which the reviewer does not subscribe. The identity between gravitational and inertial mass is stated to be one of the curiosities of classical mechanics, a curiosity which was first made clear by Einstein's general theory of relativity. This view has also been expressed by others. However, a distinction between gravitational and inertial mass is not inherent in Newtonian mechanics although it does seem to be inherent in the mode of development which regards force as a fundamental idea. The trucks shown in Fig. 46 would afford an excellent means of introducing the mass concept without referring to forces an without introducing such vague terms as inertia or heaviness to account for it. One of the dangers in the use of force as a fundamental idea is that of attributing to it a "causal" significance rather than a "descriptive" one.

J. W. Campbell
University of Alberta

CURRENT LITERATURE OF PHYSICS

Extracts from the Tables of Contents of Some of our Contemporaries

NATURE

VOL. 129, No. 3265 — MAY 28, 1932

Letters to the Editor:

Ionization at High Gas Pressures. By **R. M. Sievert** 792
Atomic Weight of Fluorine. By **H. S. Patterson and W. Cawood** 794
Vibrational Specific Heat of Carbon Dioxide. By **H. O. Kneser** 797
Types of Iridescent Clouds. By **H. S. Wynne-Edwards and T. S. Dymond** 798
Musical Sands. By **C. Carus-Wilson** 798
Neutron Hypothesis. By **D. Iwanenko** 789

NATURE

VOL. 129, No. 3266 — JUNE 4, 1932

Letters to the Editor:

Nuclear Spin of Phosphorus from the Band Spectrum. By **F. A. Jenkins and Muriel Ashley** 829
Disintegration of Atomic Nuclei. By **H. S. Allen** 830
Oscillations of the Methane Molecule. By **S. Bhagavantam** 830
Errors in Thermal Measurements. By **W. E. Garner** 832
Free Ethyl. By **T. G. Pearson, P. L. Robinson, and E. M. Stoddart** 832
Electrostatic Explanation of the Phenomenon of Flotation. By **Ian W. Wark** 833
Cry of Tin. By **Hugh O'Neill** 833
Isotopic Constitution of Lead. By **H. Schüler and E. Gwynne Jones** 833

DIE NATURWISSENSCHAFTEN

VOL. 20, Nos. 22/24 — MAY 27, 1932

Proceedings and Annual Reports of the Kaiser Wilhelm-Gesellschaft zur Förderung der Wissenschaften

New Electrochemical Processes of Radiochemistry. By **Otto Erbacher** 390
On the Theory of Atomic Disintegration by Resonance. By **H. Kallmann** 393

ANNALEN DER PHYSIK

VOL. 13, No. 4 — MAY 4, 1932

Faraday Effect in Fluid Mixtures. By **K. Scharf** 377
Measurements between Two Perfectly Conducting Planes in the Radiation Field of a Vertical Antenna Excited to its Fundamental Vibration. By **L. Bergmann and W. Doerfel** 409
Theoretical Considerations about Cosmic Radiation. By **W. Heisenberg** 430
Electron Diffraction and Molecular Structure. II. By **R. Wierl** 453
On Changes in the Striking Place of Long Electric Sparks without Change of Spark Gap. By **B. Walter** 483

COMPTES RENDUS

VOL. 194, No. 20 — MAY 17, 1932

Effect of Hysteresis in Heating by an Oscillating Magnetic Field. By **André Blondel** 1700
Diameter of CO Molecule. By **E. Mathias, W. J. Bijleveld and Ph. P. Grigg** 1708
Variation of Drag at Low Velocities under the Influence of the Compressibility. By **Basile Demtchenko** 1720

On the Role Played by the Nature of the Electrodes in the Conductivity of Liquid Semiconductors. By **J. Sambussy** 1724
On the Variable Paramagnetism of Crystallized Iron Perchloride and the Constant Paramagnetism of the $Fe^2 Cl^6$ Molecule in the Gaseous State. Constitution of the Molecule. By **André Lallemand** 1726
Double Salts, Complex Salts and Circular Dichroism. By **J. P. Mathieu** 1727
Stroboscopic Method for the Measurement of Electric Double Refraction. By **Réne Lucas and Marcel Schwob** 1729
Actinium Uranium Ratio in an Old Uraninite, the cleveite of Aust-Agder (Norway). By **Ellen Gleditsch and Sverre Klemetsen** 1731
Radioactive Phenomena of the Second Order and of Artificial Origin. By **G. Reboul** 1733
Effect of Substiuitions on the Vibration Frequencies of Ethylene Compounds. Procedure for the Classification of Radicals. By **M. Bourguel** 1736

COMPTES RENDUS

VOL. 194, NO. 21 MAY 23, 1932

On the Temperature of a Piezoelectric Crystal as a Function of its State of Oscillation. By **Armand de Gramont and Daniel Beretzki** 1777
Application of the Method of the Variation of Constants to the Problem of n Bodies with Variable Masses. By **M. Mendes** 1794
On Some Problems of Elastic Vibrations. By **V. Smirnoff and S. Soboleff** 1797
Discussion of a Nernst Bridge for High Precision. By **Jacqueline Hadamard** 1799
Influence of Solvents and of the Temperature on the Optical Activity and Rotary Dispersion of Active Substances. By **R. Lucas and D. Biquard** 1805
Role of Multiple Reflection in the Magneto-Optic Kerr Effects of Thin Sheets of Iron. By **Marcel Cau** 1807
Influence of Diffused Light on Photoelectric Measurements. By **Trajan D. Gheorghiu** 1810
Passage of α-Rays through Air and the Theory of Bethe. By **G. Mano** 1813
Transparency of the Pure Atmosphere. By **J. Duclaux and M. Hugon** 1842
Transparency of Air. By **P. Chofardet** 1844